

Universidade do Minho
Escola de Engenharia

Ana Sofia Martins Sá Quintas

**Aprendizagem não supervisionada
de padrões de Interação Homem-Computador**



Universidade do Minho
Escola de Engenharia

Ana Sofia Martins Sá Quintas

Aprendizagem não supervisionada de padrões de Interação Homem-Computador

Dissertação de Mestrado
Mestrado Integrado em Engenharia Biomédica
Ramo Informática Médica

Trabalho efetuado sob a orientação de
Paulo Novais

Agradecimentos

Não posso deixar de expressar um agradecimento especial aos meus orientadores, assim como às pessoas sem as quais este trabalho não seria o mesmo.

Ao Professor Doutor Paulo Novais agradeço as suas orientações, conselhos e disponibilização de recursos que foram fundamentais para a concretização da presente dissertação.

Ao mestre e aluno do programa doutoral em ciências da computação André Pimenta agradeço a possibilidade de trabalhar com uma das *startups* mais promissoras do nosso país. Agradeço ainda, por todas as orientações e disponibilidade demonstrada ao longo da várias etapas da dissertação.

À minha família, e em especial aos meus pais e irmã por sempre acreditarem em mim e me darem motivação diária.

Por último, um profundo agradecimento a todos os meus amigos, em especial às minhas companheiras de casa, que me ouviram e apoiaram e que permitiram, mesmo que inconscientemente ou por métodos menos convencionais, a resolução de vários dos obstáculos encontrados.

Resumo

Durante longos períodos de atividade cognitiva é comum a sensação de cansaço e falta de energia, acompanhada de um decréscimo de desempenho. Este estado, geralmente denominado de fadiga mental, é considerado uma das principais causas de erro humano. Os efeitos da fadiga mental no desempenho de tarefas complexas e que requerem altos níveis de concentração devem ser estudados e antecipados de forma a minimizar erros. Exemplo disto é o caso da aviação, ou medicina onde pequenas distrações podem gerar graves acidentes.

O impacto negativo da fadiga mental na performance, saúde e bem estar dos indivíduos, torna-se, desta forma, um dos principais motivos que leva ao desenvolvimento de metodologias de deteção destes estados mentais, de forma a preveni-los. Efetivamente, ao longo do tempo diversas metodologias de deteção de fadiga têm sido desenvolvidas, contudo a maioria é pouco objetiva ou requer um grande investimento económico.

A presente monografia apresenta um estudo de um sistema de aprendizagem não supervisionada de casos de fadiga mental, utilizando padrões de interação homem-computador, recolhidos através da sensorização de rato e teclado. Este estudo permitiu uma classificação correta de 83,3% de novos casos de fadiga mental.

Abstract

During long periods of cognitive activity it is common to have a feeling of tiredness and lack of energy, which is followed by a performance decrease. This state, commonly defined as mental fatigue, is considered one of the major causes of human error. Mental fatigue effects on tasks that require high levels of concentration should be studied and anticipated to minimize errors. Aviation and medicine are good examples of these tasks because small distractions could lead to severe consequences.

The negative impact of mental fatigue on individuals performance, health and well-being, as well as the consequences of it, makes it necessary to develop strategies that detect these mental states. During the past years many methods have been developed. However, the majority of these methods are not very objective and requires a large economic investment.

The following dissertation presents a non-invasive automatize system for unsupervised learning in mental fatigue, by using human-computer interaction patterns, collected through mouse and keyboard sensing. This study allowed a correct classification of 83.3 % of new cases of mental fatigue.

Índice

Agradecimentos	i
Resumo	iii
Abstract	v
Lista de Tabelas	ix
Lista de Figuras	xiii
Lista de acrónimos	xv
1 Introdução	1
1.1 Motivação	3
1.2 Objetivos	3
1.3 Metodologia de Investigação	4
1.4 Estrutura do Documento	5
2 Fadiga	7
2.1 Conceito de fadiga	7
2.2 Fadiga mental	10
2.2.1 Fatores da fadiga mental	10
2.3 Detecção de fadiga mental : Estado da arte	13
2.3.1 Sistemas baseados em avaliações subjetivas	13

2.3.2	Sistemas baseados em parâmetros comportamentais	14
2.3.3	Sistemas baseados em parâmetros fisiológicos	18
2.4	Conclusão	21
3	Interação Homem-Computador	23
3.1	Conceito de HCI	23
3.1.1	O Homem	24
3.1.2	O computador	25
3.1.3	A interação	26
3.2	HCI na avaliação de estados mentais	27
3.2.1	Emoções e estados afetivos	28
3.2.2	Personalidade	29
3.2.3	Fadiga Mental	29
3.3	Conclusão	29
4	Aprendizagem não supervisionada	31
4.1	Inteligência Artificial	32
4.1.1	Aprendizagem	33
4.2	Clustering	34
4.2.1	Medidas de similaridade	36
4.2.2	Técnicas de clustering	37
4.2.3	Avaliação dos clusters	44
4.3	Conclusão	46
5	Caso de Estudo	47
5.1	Metodologia	47
5.2	Preparação dos dados	48
5.2.1	Processamento dos dados	50
5.3	Pré análise dos dados	51
5.3.1	Key Down Time	54

5.3.2	Mouse acceleration	55
5.3.3	Mouse velocity	56
5.3.4	Double Click Duration	58
5.3.5	Average Excess of Distance	59
5.3.6	Average Distance of the Mouse to the Straight Line	61
5.4	Conclusão	63
6	Processo de aprendizagem	67
6.1	Análise de agrupamentos	67
6.1.1	K-means	68
6.1.2	PAM	70
6.1.3	Método hierárquico aglomerativo	71
6.1.4	Fuzzy c-means	73
6.2	Previsão de fadiga	75
6.2.1	k-means	75
6.2.2	PAM	76
6.3	Discussão de resultados	76
7	Conclusões e Trabalho Futuro	79
7.1	Considerações sobre o trabalho	80
7.2	Trabalho futuro	81
	Bibliografia	83
	Apêndices	91
A	Clustering segundo a escala USAFSAM	93

Lista de Tabelas

2.1	Tabela resumo dos trabalhos realizados no âmbito da deteção de fadiga com sistemas baseados em parâmetros comportamentais	16
2.2	Tabela resumo dos trabalhos realizados no âmbito da deteção de fadiga com sistemas baseados em parâmetros fisiológicos	18
4.1	Exemplo de matriz de confusão	45
5.1	Síntese dos dados	51
5.2	Resultado da análise de Mann-Whitney	53
5.3	Resultado da análise de Kruskal-Wallis	53
5.4	Valores médios dos dados para o KDT	55
5.5	Valores médios dos dados para o MA	55
5.6	Análise dos grupos USAFSAM para a métrica MA	56
5.7	Valores médios dos dados para o MV	57
5.8	Valores médios segundo a escala USAFSAM, relativas ao MV	58
5.9	Valores médios dos dados para o DCD	59
5.10	Valores médios dos dados para o AED	60
5.11	Valores médios segundo a escala USAFSAM, relativas ao AED	61
5.12	Valores médios dos dados obtidos para o ADMSL	62
5.13	Valores médios segundo a escala USAFSAM, relativas ao ADMSL	63

6.1	Matrizes de confusão para o algoritmo K-means, na distinção entre estados de ausência de fadiga e estados de fadiga mental	68
6.2	Médias dos agrupamentos para o algoritmo K-means, na distinção entre ausência de fadiga e estados de fadiga mental	69
6.3	Medidas de avaliação interna para o algoritmo k-means na comparação na distinção entre ausência de fadiga e estados de fadiga mental.	69
6.4	Medidas de avaliação externa para o k-means	70
6.5	Matrizes de confusão para o algoritmo PAM, na distinção entre estados de ausência de fadiga e estados de fadiga mental	70
6.6	Médias dos agrupamentos para o algoritmo PAM, na distinção entre estados de ausência de fadiga e estados de fadiga mental	71
6.7	Medidas de avaliação interna para o algoritmo PAM	71
6.8	Medidas de avaliação externa para o PAM	71
6.9	Matrizes de confusão para o algoritmo hierárquico aglomerativo, na distinção entre estados de ausência de fadiga e estados de fadiga mental	72
6.10	Medias dos agrupamentos para o algoritmo hierárquico aglomerativo, na distinção entre estados de ausência de fadiga e estados de fadiga mental	72
6.11	Medidas de avaliação interna para o algoritmo hierárquico aglomerativo	72
6.12	Medidas de avaliação externa para o método hierárquico aglomerativo	73
6.13	Matrizes de confusão para o algoritmo fuzzy c-means, na distinção entre estados de ausência de fadiga e estados de fadiga mental	73
6.14	Médias dos agrupamentos para o algoritmo fuzzy c-means, na distinção entre ausência de fadiga e estados de fadiga mental	74
6.15	Medidas de avaliação interna para o algoritmo fuzzy c-means	74
6.16	Medidas de avaliação externa para o algoritmo fuzzy c-means	74
6.17	Matriz de confusão na previsão de novos casos de fadiga para o algoritmo k-means	75
6.18	Índices de avaliação de previsão dos novos casos para o algoritmo k-means	75
6.19	Matriz de confusão na previsão de novos casos de fadiga para o algoritmo PAM	76
6.20	Índices de avaliação de previsão dos novos casos para o algoritmo PAM	76

6.21	Tabela comparativa para o modelo geral	77
6.22	Tabela comparativa para o modelo individual	77
6.23	Tabela resumo da previsão de novos casos de fadiga	78
A.1	Tabela k-means segundo a escala USAFSAM	93

Lista de Figuras

3.1	Etapas de atividades entre o golfo da execução e ação	28
4.1	Exemplo de agrupamento [68]	34
4.2	Agrupamento com dois clusters [68]	35
4.3	Agrupamento com quatro clusters [68].	35
4.4	Agrupamento com seis clusters [68]	35
4.5	Distância de manhattan e distância euclidiana para o ponto i e j	37
4.6	Exemplo de dendograma	42
5.1	Diagrama de caixa para a métrica ADMSL	50
5.2	Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao KDT	54
5.3	Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao MA	56
5.4	Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao MA	57
5.5	Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao MV	58
5.6	Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao MV	59
5.7	Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao DCD	60
5.8	Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao AED	61
5.9	Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao AED	62
5.10	Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao ADMSL	63
5.11	Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao ADMSL	64

Acrónimos

ADMSL Average distance of the mouse to the straight line

AED Average excess of distance

DCD Double click duration

DMSL Distance of the mouse to the straight line

ECG electrocardiograma

EEG electroencefalograma

EMG electromiografia

HCI Interação Homem-Computador

IA Inteligência artificial

KDT Key Down Time

MA Mouse acceleration

MV Mouse velocity

PAM Partition Around Medoids

PERCLOS Percentage of Eye Closure

TBC Time between clicks

TBK Time between keys

Capítulo 1

Introdução

A fadiga trata-se de um fenómeno complexo e de experiência pessoal que envolve um vasto leque de componentes e processos psicológicos e comportamentais. Este conceito apresenta diversas abordagens, sendo que esta monografia se foca na fadiga mental [73].

A fadiga mental trata-se de uma forma de manifestação de fadiga experienciada geralmente após longos períodos cognitivos, mas não só. Este é um problema que afeta o dia-a-dia, que se manifesta através de cansaço, perda de performance e problemas de concentração [12]. Esta é considerada o principal motivo de falha humana, sendo que por isso deve ser atempadamente solucionada [43].

Ao longo do tempo tem aumentado a preocupação em solucionar este problema, no entanto as soluções desenvolvidas apresentam um elevado grau de subjetividade ou a necessidade de utilização contínua de dispositivos que nem sempre se apresentam confortáveis ou estéticos [49]. Assim, com vista a colmatar estas limitações, é necessário o desenvolvimento de uma estratégia automatizada dotada da capacidade de identificar cenários de fadiga mental, sonolência e/ou carga de trabalho mental, de forma totalmente não invasiva. É neste ponto que a análise de padrões de Interação Homem-Computador (HCI), associados à Inteligência artificial (IA), nomeadamente à aprendizagem não supervisionada, podem-se revelar muito úteis.

O conceito de HCI surgiu em 1980 e trata-se de um campo de estudo que analisa a forma como os utilizadores

interagem com os computadores, objetivando melhorar a interação homem-máquina [21]. De forma a compreender este sistema de interação é necessário perceber como o homem se relaciona e interage, com o intuito de tornar a interação com o computador e tecnologias relacionadas, o mais natural possível.

O homem relaciona-se de uma forma muito mais profunda do que apenas uma relação física com os objetos. É sabido que o ser humano interage com o computador através das suas interfaces. Duas das mais tradicionais interfaces de comunicação homem-máquina referem-se ao rato e ao teclado [35]. O recurso a estes dispositivos com o intuito de reconhecer emoções, estados afetivos e identificação de usuários, já foi alvo de estudo no passado [37],[50, 51]. Para tal, no caso da utilização do rato, são obtidos dados como a velocidade, tempo entre cliques ou distâncias percorridas. No que se refere ao uso do teclado, medidas como o tempo de pressão das teclas, velocidade de digitação, número de erros podem ser extraídas [51, 56]. Os dados recolhidos são posteriormente processados de modo a reconhecer padrões.

Uma das formas possíveis de análise de dados e reconhecimento de padrões, utiliza o recurso à aprendizagem não supervisionada, também denominada de *clustering* [14]. Esta abordagem distingue-se da aprendizagem supervisionada na medida em que o sistema aprende de forma independente e sem a intervenção de um “supervisor”. Este modo de aprendizagem consiste na utilização dos dados de entrada do sistema, não classificados, de forma a que a rede encontre as regularidades estatísticas entre eles. Para tal ser concretizável, os dados devem apresentar redundância. Desta forma, é então possível reconhecer padrões que levam ao seu agrupamento em *clusters* [31, 14].

A aprendizagem não supervisionada possui diversas aplicações, são exemplo disso a segmentação de imagens, muito úteis em imagiologia, ou o reconhecimento de objetos. Outras aplicações com igual destaque, referem-se à sua utilização na recuperação de informação na WEB e no processo de Data Mining [31].

Com base nos conceitos apresentados, pretende-se, então, elaborar um sistema automatizado não invasivo, para aprendizagem não supervisionada de casos de fadiga, utilizando padrões de interação homem-máquina, recolhidos através da sensorização de rato e teclado.

1.1 Motivação

A fadiga mental é notória no dia-a-dia, após longos períodos de grande esforço mental, tratando-se de um dos grandes problemas da atualidade [73]. Esta pode ser vista como um fator que afeta negativamente o desempenho dos trabalhadores e é ainda considerada um dos principais motivos de erro humano [43]. Exemplo disto é o caso da aviação, ou medicina onde pequenas distrações podem gerar graves acidentes [76].

A influência da fadiga mental no desempenho de tarefas diárias, torna-se, desta forma, um dos principais motivos que leva ao desenvolvimento de metodologias de detecção destes estados mentais, de forma a preveni-los. Hoje, estas abordagens estão geralmente relacionadas com o uso de questionários e mais recentemente recorrendo a sensores. No entanto, se por um lado a primeira abordagem torna-se muito subjetiva e pouco precisa na medida em que é difícil ter consciência do seu próprio estado, por outro lado, é notório que nem sempre os utilizadores estão disponíveis para responder a questionários, e a fazê-lo da forma mais honesta possível. Relativamente à segunda abordagem, apesar de bastante mais precisa, o investimento económico necessário ou a necessidade de recorrer a aparelhos estranhos ao quotidiano, por vezes incómodos, pode-se tornar uma grande desvantagem. [51]

Esta dissertação é então motivada pela necessidade de colmatar as limitações anteriormente referidas. Uma vez que nos dias de hoje muitos profissionais passam grande parte do seu tempo, ou a totalidade do mesmo, em frente a um computador, parte-se do princípio que se poderá adaptar esta ferramenta com o qual já estão ambientados para proceder à detecção de fadiga, sem que com isto traga incómodos ou grandes custos. Este tipo de ferramenta tornar-se-ia útil não só para uma melhoria da qualidade de vida, assim como um aumento da produtividade, mas também contribuiria para a existência de menos erros associados à fadiga.

1.2 Objetivos

O objetivo principal desta dissertação trata-se da utilização de algoritmos não supervisionados para aprendizagem de padrões de HCI com o intuito de identificar cenários de fadiga mental, sonolência ou carga de trabalho mental, através da sensorização do rato e teclado (usando biométricas comportamentais) de forma totalmente não invasiva. De seguida enumera-se de forma mais específica os objetivos do trabalho proposto:

- Revisão bibliográfica da área, de forma a caracterizar o domínio de investigação;
- Avaliar o desempenho das diferentes abordagens e a aplicabilidade no domínio de investigação;
- Desenvolvimento de uma abordagem não supervisionada para detecção de fadiga mental, sonolência e ou carga de trabalho mental.
- Teste e validação do modelo desenvolvido.
- Comparação entre modelos supervisionados e não supervisionados;

1.3 Metodologia de Investigação

Na realização desta monografia foi adotada uma metodologia ação-pesquisa. Esta metodologia consiste na identificação e caracterização do problema de forma a que possa posteriormente ser resolvido.

Desta forma, efetuou-se uma coleta de informação com relevância para o problema em questão, posteriormente procedeu-se a uma organização da mesma. Foi necessário ter em consideração a constante evolução sentida na área de forma a acompanhar a mudança do conteúdo científico constante. Após a correta identificação do problema em causa foi concebida uma hipótese de resolução. Após a validação desta hipótese procedeu-se ao seu desenvolvimento e aplicação. Por último foram tiradas conclusões de forma a avaliar os resultados obtidos. Resumidamente o projeto passou pelas fases de:

1. Identificação do problema em questão;
2. Coleta de informação relativa ao estado da arte;
3. Desenvolvimento de uma solução e posterior implementação;
4. Análise e discussão dos resultados;
5. Validação da solução.

1.4 Estrutura do Documento

O presente documento apresenta-se organizado da seguinte forma: no Capítulo 2 apresenta-se o conceito de fadiga e mais especificamente o conceito de Fadiga Mental, problemas associados e revisão dos estudos e sistemas de deteção já existentes e implementados. O Capítulo 3 aborda a definição de HCI e principais requisitos da interação homem-computador. No Capítulo 4 apresenta-se o conceito de inteligência artificial e aprendizagem, com destaque para aprendizagem não supervisionada e algoritmos de *clustering* que serão utilizados no Capítulo 6. No Capítulo 5 apresenta-se o caso de estudo, e uma análise detalhada do conjunto de dados cedido pela *startup Performetric*. No Capítulo 6 apresenta-se a aplicação dos algoritmos não supervisionados, comparações entre os demais e respetivas conclusões. O documento é finalizado com o Capítulo 7 onde estão descritas as principais conclusões sobre o trabalho efetuado, assim como o futuro trabalho planeado.

Capítulo 2

Fadiga

De uma forma geral, a fadiga pode ser definida como um estado que engloba dimensões físicas e/ou mentais e que resulta num impacto negativo na performance e processamento de informações. Estas dimensões estão estritamente relacionadas e não devem ser tratadas de forma independente. Este estado, é considerado uma das principais causas de erro humano, sendo que o seu impacto já foi estudado em diversas áreas, desde aviação, medicina, entre outras. Para além do impacto imediato, a fadiga também pode ter consequências a longo prazo na saúde mental e física dos indivíduos.

2.1 Conceito de fadiga

A fadiga trata-se de um fenómeno complexo e subjetivo, que envolve uma série de processos psicossociais e comportamentais. Como resultado, ao longo do tempo, diversos conceitos têm sido apresentados, sendo que não existe um unicamente aceite [62, 17]. Uma das possíveis definições apresenta-se como *"... uma enorme sensação de cansaço, perda de energia e exaustão, associada de uma diminuição do desempenho físico e do funcionamento cognitivo ..."* [62].

Apesar da definição anteriormente apresentada, existem diversas linhas literárias que fazem a distinção entre diferentes conceitos de fadiga, recorrendo a uma dualidade entre definições. As mais comuns distinguem fadiga aguda de fadiga crónica, fadiga física de fadiga mental e ainda fadiga central de fadiga periférica [62]. Estas distinções serão

abordadas mais pormenorizadamente nas seguintes secções.

2.1.0.1 Fadiga crónica vs fadiga aguda

A fadiga crónica, também designada por síndrome da fadiga crónica (SFC), caracteriza-se por uma sensação inexplicável de fadiga que se prolonga há pelo menos seis meses, que não melhora após o repouso, e que afeta os níveis de atividade social, pessoal, educacional e ocupacional do indivíduo. Associados a estes sintomas, devem, ainda, ser detectáveis pelo menos quatro dos seguintes sinais: perturbações de memória e concentração, sono não revitalizante, dores musculares, dores nas articulações sem inchaço ou vermelhidão aparente, ganglios linfáticos cervicais e axilares "moles", dores de cabeça diferentes do habitual, mal estar após esforço que dura mais de 24 horas e dores de garganta frequentes [17, 42]. O diagnóstico da SFC exige atenção ao historial clínico e o despiste de um conjunto de doenças que se podem assemelhar nos sintomas, nomeadamente doenças relacionadas com a tiroide, anemia e depressão [42].

A fadiga aguda, por outro lado, surge em indivíduos saudáveis como resultado de uma defesa natural do corpo. Esta manifesta-se após algum tipo de esforço e que desaparece após o descanso. O seu impacto no dia-a-dia é mínimo [62].

Assim sendo, as principais distinções entre os conceitos anteriormente apresentados assentam na duração da manifestação da fadiga e nas consequências que apresentam. Isto porque, quanto mais se prolongar o estado de fadiga, não só maior é a gravidade da situação, mas também implica um tratamento mais complexo.

2.1.0.2 Fadiga física vs fadiga mental

A fadiga física compreende uma diminuição na performance física e o aumento da dificuldade de cumprir um exercício ou atividade. De um ponto de vista muscular, trata-se da incapacidade para manter o nível necessário de força durante os exercícios. Esta denota-se ao fim de algum tempo de esforço físico e pode ser detetável através de electromiografia [2, 62, 43].

A fadiga mental, por sua vez, trata-se de um estado complexo e multifacetado que envolve mudanças na motivação, humor e nível cognitivo. Este tipo de fadiga tem um profundo impacto na performance mental e é considerado

um das principais causas de erro e acidente [28]. A fadiga mental pode ser experienciada após longos períodos de atividade cognitiva, onde os indivíduos apresentam uma menor capacidade de manter concentração e processamento de informação. Isto resulta num menor envolvimento com a tarefa e numa diminuição da capacidade de resposta e de ignorar informações irrelevantes [10, 12]. No entanto, estudos mostraram que nem sempre um maior número de horas de atividade leva necessariamente a fadiga. Para além do tempo, esta envolve também uma componente motivacional para do indivíduo no cumprimento da mesma. Por outras palavras, uma alta carga de trabalho pode resultar em fadiga, se associado estiver uma pequena recompensa [11].

Apesar de dois conceitos distinguíveis, a fadiga mental e fadiga física, estão intrinsecamente relacionadas e fortemente dependentes. Esta relação suporta a existência de um outro conceito de fadiga: a fadiga total. A fadiga total pode ser vista como a associação de duas dimensões: a fadiga mental e a fadiga física, associadas a perturbações do sono, stress, e fatores psicológicos como o humor [6].

2.1.0.3 Fadiga central vs fadiga periférica

A origem da fadiga muscular pode dever-se a causas relacionadas com o cérebro (mecanismos centrais), assim como causas relacionadas com os próprios músculos (mecanismos periféricos) [2].

A fadiga central, apesar de ainda pouco explorada e com mecanismo que ainda não foram completamente desvendados, relaciona-se com o aumento da concentração de serotonina, um neurotransmissor, que se acumula em certas regiões cerebrais após períodos de atividade física prolongados. Isto resulta numa falha da condução do impulso nervoso, seguindo-se uma diminuição do número de unidades motoras recrutadas para a execução do movimento, e consequentemente um decréscimo da performance [2].

Por outro lado, a fadiga periférica caracteriza-se por fenómenos que decorrem do exercício a nível muscular, que levam a uma diminuição da força de contração das fibras musculares e alteração dos mecanismos de transmissão do potencial de ação. As alterações anteriormente referidas são principalmente devidas a perturbações do movimento de iões de cálcio, acumulação de fosfato e diminuição da adenosina trifosfato (ATP) disponível [2].

2.2 Fadiga mental

A fadiga mental, como já foi referido anteriormente, trata-se de um tipo de fadiga bastante comum no dia a dia, que se relaciona com a incapacidade de cumprir determinada tarefa cognitiva [10].

Durante longos períodos de atividade cognitiva é comum a sensação de cansaço e falta de energia, acompanhada de um decréscimo de desempenho [10, 12, 43]. Este estado, é considerado uma das principais causas de erro humano [76]. Os efeitos da fadiga mental no desempenho de tarefas complexas e que requerem altos níveis de concentração devem ser estudados e antecipados de forma a minimizar erros. Exemplo disto é o caso da aviação, ou medicina onde pequenas distrações podem gerar graves acidentes [10, 73].

A influência da fadiga mental no desempenho de tarefas diárias, torna-se, desta forma, um dos principais motivos que leva ao desenvolvimento de metodologias de deteção e monitorização destes estados mentais, de forma a preveni-los.

2.2.1 Fatores da fadiga mental

A fadiga mental é definida como um estado manifestado após longos períodos de atividade cognitiva [68]. Manifesta-se como uma redução na performance do indivíduo. No entanto, estes não são afetados de igual forma. Ao longo do tempo, estudos demonstraram uma relação entre a fadiga mental e diversos parâmetros comportamentais, sociais e psicológicos e do perfil individual que são influenciadores e influenciados por estes estados mentais. Estes parâmetros serão, então abordados de seguida.

2.2.1.1 Performance mental

A fadiga mental afeta de forma crítica a performance mental numa variedade de atividades. Tal como já mencionado, a fadiga mental manifesta-se através de um decréscimo de desempenho durante o cumprimento de determinada tarefa, sendo que as mais complexas são mais afetadas do que as simples [65]. Assim, a performance mental é caracterizada pela incapacidade de alocar os recursos necessários para a conveniente realização das tarefas que se está a desempenhar [64, 40]. Esta relaciona-se com algumas capacidades cognitivas como o tempo de reação, a memória,

a atenção e a precisão. Sendo que o conjunto destes parâmetros avaliam o nível de performance mental do indivíduo. De seguida estes parâmetros serão discutidos individualmente.

- **Memória de trabalho:** A memória de trabalho pode ser definida como a capacidade cerebral de recolher, organizar, armazenar e recuperar informações. Esta é um alicerce fundamental na aprendizagem e tomadas de decisões. Uma das formas de manifestação de fadiga mental trata-se da diminuição da capacidade de armazenamento de novas informações, assim como dificuldades na recuperação das mesmas [71, 32].
- **Reação:** A reação caracteriza-se como um comportamento manifestado em relação a um determinado estímulo. Este comportamento pode manifestar-se de várias formas quer através de um movimento ou alteração de sinais vitais. A associação da fadiga mental com o cansaço resultam num aumento do tempo de reação e consequentemente numa menor capacidade de reação [11].
- **Atenção:** Quando as pessoas se sentem fatigadas reportam queixas relacionadas com dificuldades em manter o foco na tarefa que estão a desempenhar. A atenção não só influencia o processamento das informações recolhidas como auxilia na concentração daquilo que é essencial, desviando o interesse de informações ou situações irrelevantes. Assim, em termos de fadiga mental, os indivíduos apresentam dificuldades em manter a atenção durante a execução das tarefas que lhe são apresentadas [10, 11].
- **Precisão:** A capacidade de obter resultados dentro do expectável denomina-se de precisão. Esta capacidade que pode ser expressa através do número de erros obtidos e encontra-se relacionada com a fadiga mental. Assim, quando estamos perante de um quadro de fadiga mental, existe uma maior probabilidade de se cometerem erros, logo a precisão é menor [10].

2.2.1.2 Perfil do indivíduo

O perfil individual de cada um é determinante no processo de deteção de fadiga. Fatores como a idade e o sexo e doenças associadas deverão ser tidos em conta [24].

- **Idade:** A idade é um fator preponderante que influencia a fadiga mental. Como já foi referido com o avançar da idade, as capacidades cognitivas e físicas vão se deteriorando o que se manifesta num maior frequência de sensação de fadiga [24] .

- Neuropatologias: A fadiga mental está muitas vezes associada a doenças do campo neuronal. Em quadros clínicos de Parkinson, depressão, esclerose múltipla, entre outras, os pacientes apresentam queixas relativas a uma sensação de cansaço e fadiga muito mais precocemente do que quando comparados com indivíduos saudáveis [16, 24].

2.2.1.3 Outros

Para além dos fatores já apresentados, existem ainda um conjunto de parâmetros que não tendo uma relação direta com a fadiga mental, devem ser tidos em conta de forma a caracterizar melhor o domínio da fadiga mental. Estes devem ser considerados como potenciais causadores destes estados, como é o caso da sonolência, motivação, stress e má alimentação. Estes apresentam-se discutidos de seguida.

- Sonolência: A sonolência é descrita como sinal que antecede o sono. Esta apresenta-se muitas vezes relacionada com o conceito de fadiga mental, sendo difícil distingui-los. Facto é que a sonolência é um dos principais indicadores de fadiga, e esta por sua vez, encontra-se relacionada com queixas de sono não reparadores [5, 47].
- Motivação: A fadiga pode ser vista como um desequilíbrio entre o esforço e a recompensa. Um indivíduo manterá o seu nível de desempenho se o esforço que está a aplicar no desempenho da tarefa for recompensado. No entanto quando o esforço é demasiado grande e a recompensa já não acompanha o mesmo nível, a motivação desvanece e o indivíduo sente-se fatigado mais rapidamente [11].
- Stress: O stress relaciona-se com a ansiedade, excesso de trabalho, medo e preocupação. O stress pode potenciar a manifestação de sintomas de fadiga mental, além disso poderá ser a causa, ou potenciar, falta de atenção aos detalhes, más decisões e erros não intencionais [23].
- Alimentação: Tal como já referido, uma das manifestações de fadiga mental é a sensação de perda de energia. Assim, esta sensação pode relacionar-se com uma alimentação pobre em glucose. Uma alimentação cuidada pode ajudar a prevenir erros e distrações assim como manter o nível de produtividade, evitando desta forma a sensação de fadiga [23].

2.3 Detecção de fadiga mental : Estado da arte

Uma vez que o conceito de fadiga mental é complexo e extenso, ao longo do tempo têm surgido diversas formas de deteção destes estados, que variam consoante o seu propósito. A forma como é feita esta deteção divide-se essencialmente em três: de forma subjetiva recorrendo maioritariamente a questionários, através da análise de comportamentos (expressões faciais, piscar de olhos, movimentos da boca entre outros) e por fim através de sinais fisiológicos [55]. Atualmente grande parte dos sistemas de deteção de fadiga mental têm a sua aplicação na viação, no entanto já existem outros sistemas com diferentes aplicações. De seguida encontra-se uma análise de cada um dos sistemas.

2.3.1 Sistemas baseados em avaliações subjetivas

Como já referido, as escalas subjetivas avaliam o nível de fadiga através de um questionário, onde são colocadas questões e o indivíduo responde a cada item colocado dentro de um intervalo. Posteriormente, as respostas são avaliadas e chega-se a uma quantificação da fadiga mental dentro de uma escala. Estes questionários, na sua maioria não só abordam a deteção de fadiga mental, mas também incluem outras dimensões como é o caso da fadiga crónica e fadiga física. O facto de se tratar do indivíduo a auto-classificar introduz um grande nível pessoal de subjetividade [48, 29]. Para além disso, não permitem uma monitorização ao longo do tempo, apenas uma avaliação pontual.

De seguida apresenta-se algumas das escalas de fadiga mais utilizadas.

- Visual Analog Scale for Fatigue (VAS-F): Esta trata-se de uma escala destinada a adultos entre os 18 e 55 anos composta por 18 itens de avaliação da fadiga e níveis de energia, tanto em indivíduos saudáveis como em paciente que apresentam falta de sono. Em cada item, o indivíduo deve classificar-se com um "X", numa escala visual de 0 a 10 que representam dois extremos (por exemplo: "nada cansado" e "extremamente cansado"). A vantagem desta escala está no facto de não exigir muito esforço de leitura, ser extremamente rápida de responder [4].
- The Chalder Fatigue scale (CFQ 11): O propósito desta escala é a classificação da fadiga em geral, sendo composta por 11 níveis. Para obtenção desta classificação os indivíduos deverão responder a um questionário dividido em dois grandes grupos de questões: questões relacionadas com fadiga mental e questões relacionadas com fadiga física. No primeiro tema de questões são colocadas perguntas como: "Tem problemas de concentração?", "Como está a sua memória?". No segundo tema, são colocadas questões como: "Sente-se

fraco?”, “Precisa de descansar mais?”. Cada uma das questões é respondida dentro de um intervalo de quatro níveis, em que o mais baixo significa “Não mais que o habitual”, e o ponto mais alto representa “Mais do que o habitual”[30].

- USAFSAM Mental Fatigue Scale: Esta escala, desenvolvida em 1982 pela *School of Aerospace Medicine* visa a avaliação de fadiga experienciada durante voos. A partir da análise de fadiga e da performance manifestada durante turnos exaustivos, foi possível o desenvolvimento de um sistema de quantificação destes estados. Esta escala compreende 7 níveis, e cada nível descreve um estado:

1. Completamente alerta, acordado e responsivo;
2. Alerta, mas não na sua totalidade;
3. Níveis normais de resposta e alerta;
4. Cansado e pouco responsivo;
5. Moderadamente cansado;
6. Extremamente cansado;
7. Completamente exausto. Sem possibilidade de cumprir a função.

Esta escala tenta não só quantificar o nível de fadiga, mas a forma como esta afeta a performance do trabalhador [57].

2.3.2 Sistemas baseados em parâmetros comportamentais

A detecção da fadiga mental através da análise de comportamentos são geralmente baseados no pressuposto de que a fadiga se relaciona com o cansaço e sonolência, sendo que esta ser detetada através da análise de mudanças na face e olhos. Nestes casos, os indivíduos apresentam níveis de alerta reduzidos, evidenciam movimentos lentos das pálpebras, baixo nível de abertura dos olhos (ou até mesmo olhos fechados), bocejos e expressões faciais lentas [34]. Dentro dos comportamentos apresentados, o mais bem estudado trata-se da abertura dos olhos, que pode ser medido através do Percentage of Eye Closure (PERCLOS). O PERCLOS calcula a proporção de tempo em que os olhos estão fechados 80 % ou mais de um intervalo temporal especificado [75].

Estes sistemas, na sua maioria, recorrem a câmeras implementadas nos carros que permitem a monitorização do condutor [34]. No entanto a maior limitação do uso deste tipo de câmeras prende-se na iluminação. Estas geralmente apresentam má qualidade em cenários pouco iluminados [8]. Assim, de forma a ultrapassar esta questão foi estudada a possibilidade de utilizar um diodo emissor de luz infravermelho (LED) [8]. No entanto, apesar das melhorias introduzidas em cenários de pouca luz, durante o dia o mesmo não ocorre. Quando comparadas com câmeras anteriores apresentam pior desempenho [26]. Assim, uma das opções encontradas foi a utilização de uma câmera convencional durante o dia, e uma câmera de infravermelhos durante a noite [63, 19].

Um dos grandes desafios da deteção da utilização de vídeo na deteção de fadiga, trata-se do seu posterior processamento. Assim, depois de recolhidas, as imagens, são em tempo real processadas, e aplicados filtros, de forma a detetar o rosto e olhos e boca do condutor [8]. Posteriormente, os comportamentos registados são geralmente classificados em dois ou três níveis de fadiga ou cansaço. Nesta classificação das imagens são geralmente usados acSVM e redes neuronais. Apesar dos bons resultados obtidos, em alguns casos perto dos 100%, há necessidade de ressaltar, que alguns destes estudos foram conduzidos em ambiente de simulação e que não foram estudados indivíduos que utilizam óculos [8].

De forma a estender a deteção de fadiga mental a outras áreas, em 2013 surgiram os primeiros estudos que relacionam a fadiga com a utilização do rato e teclado. Nestes estudos foram recolhidas métricas como a velocidade do rato, tempo de pressão de teclas, duração de cliques entre outros. Estes foram posteriormente processados de forma a obter 5 níveis de fadiga mental. Para isso foi utilizado o algoritmo do vizinho mais proximo (KNN) como algoritmo supervisionado para reconhecimento dos padrões [51]. Este estudo foi evoluindo, algumas das métricas utilizadas foram eliminadas ou substituídas por outras com melhor relação com a fadiga mental. Recentemente, em 2015, foi publicado um novo estudo. Neste estudo, os níveis de deteção de fadiga mental foram enquadrados segundo a escala de fadiga USAFSAM. Na base da deteção de fadiga mental, encontra-se uma rede neuronal capaz de aumentar a precisão da deteção quando comparado com estudos anteriores.

Os estudos apresentados encontram-se esquematizados na tabela 2.1. Alguns dos estudos apresentados, evoluíram e tornaram-se produtos comercializados. De seguida, apresentam-se alguns dos desses produtos que estão no mercado [33].

Tabela 2.1: Tabela resumo dos trabalhos realizados no âmbito da detecção de fadiga com sistemas baseados em parâmetros comportamentais

Referência	Ano	Parâmetros	Método de recolha	Classificador	Precisão
[25]	2001	- PERCLOS	Câmera	Rede neuronal não paramétrica	-
[8]	2006	-PERCLOS -Duração do fecho dos olhos -Frequência do pestanejo Posição da face	Câmera e infravermelhos	Classificador de fuzzy	99,9%
[74]	2007	-Pestanejar -Bocejo, -Movimento de cabeça	Câmera de vídeo digital	SVM	96%
[59]	2008	- Boca - Bocejo	Câmera	SVM	80%
[19]	2010	PERCLOS	Duas câmeras (uma para a noite , outra no dia)	SVM	[91%-98%]
[63]	2012	Deteção da pupila	Micro câmara com infravermelhos	Rácio entre olhos abertos e olhos fechados	92%
[51]	2013	Interações através de rato e teclado	Computador (rato e teclado)	KNN	85%
[50]	2015	Interações através de rato e teclado	Computador (rato e teclado)	Rede neuronal supervisionada	81%

- Onguard

Este trata-se de um dos primeiros sistemas apresentados para detecção de fadiga em condução. O Onguard funciona com base na abertura dos olhos, sendo que emite um feixe de luz infravermelha no olho e mede a luz refletida. Este sistema, implementado em óculos faz soar um alarme quando o tempo de fecho dos olhos é superior a 0.5 segundos. A principal desvantagem deste sistema apresentado trata-se da necessidade de utilização de uns óculos para implementação do *hardware*, *Haworth 1991*

- Driver Fatigue Monitor (DFM)

O DD850 Driver Fatigue Monitor trata-se de um dispositivo desenvolvido pela empresa *Attention Technology, Inc* para detecção da fadiga com base no PERCLOS. Este sistema utiliza uma câmara instalada no veículo e luz infravermelha para identificação por contraste da pupila com o resto do rosto. De seguida, com base na medida PERCLOS faz soar um alarme quando as percentagens detetadas sejam consideradas inseguras. O DFM

apresenta algumas desvantagens e limitações sobretudo no que toca ao seu funcionamento. Este dispositivo, encontra-se limitado à sua utilização em ambientes bem iluminados e o seu utilizador não deve usar óculos [75].

- Driver State Monitor

Em 2005 a DELPHI apresentou o *Driver State Monitor*, este sistema trata-se de um avanço em relação ao anterior uma vez que utiliza uma câmara central e duas fontes de iluminação ultravioleta. Este sistema não só deteta a fadiga através do fecho dos olhos, mas também o nível de distração do condutor através da posição da cabeça. As principais melhorias verificadas encontram-se na utilização em diferentes iluminações e com utilização de óculos [22].

- faceLAB

Apresentado em 2009 pela Seeing Machines, trata-se de um sistema que têm sido utilizado em grande escala tanto em automóveis, como em camiões, aviões e submarinos. O faceLAB faz um rastreamento da cabeça recolhendo informações relativas ao movimento dos olhos, abertura da pálpebra e tamanho da pupila. Para além disto, este sistema funciona mesmo com o uso de óculos, lentes ou óculos de sol [60].

- NapZapper

O NapZapper consiste num pequeno aparelho que se coloca numa orelha e que soa um alarme quando a cabeça começa a inclinar-se e atinge um determinado ângulo. A principal desvantagem deste dispositivo trata-se do facto de não atuar de forma preventiva na deteção da fadiga e sonolência ¹.

- Performetric

A Performetric trata-se de um software de monitorização e deteção de fadiga mental. Para isso usa o rato e teclado como sensores e um sistema que recorre à inteligência artificial para processamento dos dados. Para além de ser um sistema não invasivo e não intrusivo, não requer dados extras e adapta-se a qualquer tipo de utilizador. No entanto necessita de um período curto de aprendizagem ².

¹<http://www.napzapper.com>

²<https://performetric.net>

2.3.3 Sistemas baseados em parâmetros fisiológicos

Existem uma série de parâmetros fisiológicos que podem ser associados com a fadiga. Destes parâmetros destaca-se a atividade cerebral detetada através do electroencefalograma (EEG), a atividade electródermica medida através de electromiografia (EMG) ou a frequência cardíaca detetada através de electrocardiograma (ECG), em alguns casos também se recorre à medição dos movimentos oculares através de electro-oculograma (EoG) [45].

A deteção de fadiga com base em parâmetros fisiológicos trata-se da forma de deteção mais precisa e confiável. No entanto a medição de sinais fisiológicos encontra-se muito propensa à introdução de ruído e artefactos. Isto deve-se ao movimento natural do corpo no desempenho das tarefas do dia-a-dia [44, 55]. De forma a eliminar o ruído têm sido utilizadas diversas abordagens de pré-processamento, dos quais se destacam o uso de filtros passa-baixo. Uma série de características estatísticas são então extraídas do sinal processado usando várias técnicas de extração de características, incluindo transformadas discretas de Wavelet e transformadas de Fourier [78, 39]. Posteriormente as características extraídas são classificadas recorrendo a redes neuronais, SVM, análises discriminantes lineares (LDA) e ainda outros métodos [39, 41, 58, 77].

No entanto a sua natureza muitas vezes intrusiva para a medição dos parâmetros fisiológicos, associado ao custo dos dispositivos afasta a sua utilização. De forma a contornar esta situação, têm sido utilizados dispositivos sem fio para medição de sinais fisiológicos de maneira menos intrusiva. Desta forma, são colocados elétrodos no corpo, sendo que os sinais são enviados via *Bluetooth* para o telemóvel ou outro dispositivo. Existem ainda estudos onde se optou pela utilização integrados nos bancos, volantes. Apesar disso, esta solução apresenta uma diminuição da precisão, uma vez que estão mais sujeitos a artefactos e erros de contacto [55].

Estes estudos encontram-se resumidos na tabela 2.2

Tabela 2.2: Tabela resumo dos trabalhos realizados no âmbito da deteção de fadiga com sistemas baseados em parâmetros fisiológicos

Referência	Ano	Sensor	Classificador	Precisão
[77]	2008	EEG	SVM	99.3%
[39]	2011	EEG, ECG, EoG	LDA,SVM,KNN	95-97%
[41]	2013	Variação do ritmo cardíaco	SVM	95%
[58]	2016	ECG	LDA	—

2.3.3.1 Sistemas de detecção de fadiga baseados na atividade cerebral

Os sistemas de detecção de fadiga baseados na atividade cerebral recorrem à utilização do EEG. O funcionamento destes sistemas, geralmente depende de electrodos acoplados ao couro cabeludo. O EEG permite identificar quatro tipos de ondas sinusoidais: a alfa, beta, teta e delta. As ondas delta encontram-se até aos 4 Hz e são características do sono profundo. As ondas teta encontram-se entre os 4Hz e 8 Hz e estão associados com estados de sonolência. Por sua vez, as ondas alfa, entre os 8Hz e 13 Hz, caracterizam estados mentais de relaxamento. Por último, as ondas beta, acima dos 13Hz, ocorrem no estado normal de vigília [6, 69]. Desta forma, o EEG permite a identificação de estados mentais associados a sonolência e fadiga com bastante eficácia. Nestes casos são observáveis aumento na proporção das ondas teta e alfa e uma diminuição de ondas beta [6, 1, 69].

Com base nestes princípios vários sistemas de detecção têm sido propostos. No entanto, estes sistemas podem ser considerados um pouco intrusivos, uma vez que exigem a utilização de um acessório acoplado à cabeça. De seguida seguem-se dois exemplos de sistemas já existentes e comercializados.

- SMARTCAP

O SMARTCAP trata-se de um sistema implementado num chapéu, gorro ou capacete para monitorização da fadiga destinado essencialmente a condutores e operadores. Este sistema possui a capacidade de ler a atividade elétrica cerebral através do uso de EEG. Os dados são enviados via *Bluetooth* desde o dispositivo instalado no chapéu para monitores instalados dentro do veículo. Emite um alarme quando é detetada fadiga ³.

- U-WAKE

O U-WAKE trata-se de um sistema destinado fundamentalmente a condutores, composto por uma banda colocada em volta da cabeça para recolha de EEG. Os dados recolhidos são posteriormente enviados via *Bluetooth* para a aplicação móvel associada ao dispositivo. A aplicação processa, analisa e classifica os níveis de fadiga, exibindo uma pontuação de 0 a 100. Quando a classificação se encontra acima de 80 é emitido um alarme de aviso⁴.

³<http://smartcaptech.com>

⁴<http://u-wake.com/en.html>

2.3.3.2 Sistemas de detecção de fadiga baseados em atividade eletrodérmica

A condutância de corrente elétrica na pele, também designada de atividade eletrodérmica, refere-se à facilidade com que com uma corrente elétrica é conduzida na pele. Esta altera-se em resposta à ativação do sistema nervoso simpático. Vários estudos têm comprovado a relação entre a atividade electrodérmica e o estado emocional dos pacientes, assim como com o seu estado mental, nomeadamente estados de fadiga.

De seguida apresentam-se dois exemplos da aplicação deste parâmetro fisiológica na detecção de fadiga. À semelhança da detecção de fadiga com EEG, a detecção de fadiga através da atividade eletrodérmica exige o uso continuado de dispositivos em contacto com o corpo.

- StopSleep

O StopSleep trata-se de um dispositivo que se usa como se fosse um anel. Este sistema têm como principal objetivo a detecção de fadiga e sonolência em condutores. Para tal, utiliza oito sensores de forma a detetar a atividade electrodérmica. Quando o sistema deteta sonolência ou redução de reação, avisa o utilizador. Este aviso pode ser feito através de um sinal vibratório, ou através de um sinal misto de som e vibração que depende do nível de perigo. O seu público alvo tratam-se dos condutores.⁵

- Sleepman

O Sleepman encontra-se neste momento em *crowdfunding* e trata-se de um dispositivo com formato de relógio que mede a atividade electrodermal para detecção não só de fadiga, mas também para análise do sono ajudando o utilizador a acordar na fase de sono mais correta.⁶

2.3.3.3 Sistemas de detecção de fadiga baseados na frequência cardíaca

A frequência cardíaca, detetado através de um eletrocardiograma, trata-se do número de batimentos do coração por um intervalo de tempo, geralmente minutos. Esta é regulada pelo sistema nervoso autónomo. Alguns estudos relacionam a frequência cardíaca com a fadiga mental, aferindo que esta leva a diminuição dos batimentos cardíacos por minutos registados. Desta forma é então possível a sua detecção [41].

⁵<http://www.stopsleep.biz>

⁶<http://sleeptrackers.io/sleepman-wearable-kickstarter/>

2.4 Conclusão

Ao longo deste capítulo foi apresentado o conceito de fadiga que como se verificou que pode ser abordado de diversas perspetivas. No entanto o foco desta monografia é a fadiga mental. Este apresenta-se um conceito subjetivo, pessoal e complexo que se manifesta no dia-a-dia e que muitas vezes é ignorado. No entanto é necessária a sua antecipação já que este pode ser a causa de falhas, falta de concentração e diminuição de performance.

Apesar da subjetividade inerente ao conceito de fadiga mental, várias formas de deteção têm sido desenvolvidas. De uma forma geral, as técnicas de deteção destes estados mentais podem ser divididos em três formas distintas:

- Forma subjetiva: inclui questionários, que posteriormente podem ser convertidos em escalas de medição de fadiga. Estes métodos apresentam como principal desvantagem o grau de subjetividade introduzido e não permitem uma monitorização contínua.
- Sistemas baseados em parâmetros comportamentais: Neste caso em concreto recorre-se a comportamento e baseiam-se no facto que a fadiga mental altera as reações face ao meio envolvente. Nestes sistemas destaca-se a utilização de medidas relacionadas com as mudanças na face e olhos.
- Sistemas baseados em parâmetros fisiológicos: Estes sistemas apresentam-se como sendo os mais precisos uma vez que medem alterações relacionadas com a frequência cardíaca, atividade cerebral ou atividade eletrodérmica. No entanto, são na sua maioria os mais difíceis de usar, uma vez que necessitam de sensores acoplados ao corpo. Além disso, necessitam de o investimento económico pode ser elevado.

Apesar dos estudos já realizados na área da deteção de fadiga, pode-se afirmar que a esmagadora maioria das soluções se encontra muito centrado no sector automóvel. Pode-se então afirmar que a proposta presente nesta dissertação é inovadora e permite uma monitorização não invasiva, objetiva da fadiga mental, com aplicação em ambientes informatizados, onde se recorre à utilização de um computador.

Capítulo 3

Interação Homem-Computador

A HCI aborda o estudo da relação do Homem com o computador. Esta relação, ou interação, abrange uma série de áreas tais como design, ergonomia, psicologia, sociologia, ciências da computação e afins. Esta interação ocorre através do hardware e software que compõem a interface Homem-computador.

A tecnologia deve ser usada sempre para maximizar nossas habilidades, e o uso de computadores deve ser o mais simples, seguro e agradável possível. A criação de sistemas difíceis de usar pode inviabilizar o sucesso de softwares que poderiam ser bastante úteis [21]. É neste campo que atua a HCI, objetivando a melhoria dos sistemas de forma a torna-los mais próximos dos seres humanos.

3.1 Conceito de HCI

O avanço tecnológico permitiu que o recurso a computadores pessoais se tornasse algo cada vez mais comum e enraizado na nossa cultura e dia-a-dia. É difícil imaginar um contexto, profissional ou não em que o computador não esteja presente, de forma directa ou indirecta. Consequentemente, o tempo dispensado na utilização destes objetos também aumentou, sendo que em muitas situações se torna superior ao tempo dispensado na interação com os colegas de trabalho, família ou amigos. Esta interação desenvolveu-se e aprofundou-se de tal forma que deu origem a um novo campo de estudo, a HCI [15, 21].

Assim, este conceito nasceu em 1980, impulsionado pela expansão dos computadores e o seu uso em massa [20]. A HCI apresenta-se como um campo de estudo que relaciona uma série de áreas desde o design, ciências da computação, ergonomia e sociologia e têm como objeto de estudo a forma como os usuários interagem com os computadores de forma a simplificar essa relação. Desta forma, almeja-se uma adaptação dos computadores, quer ao nível de software, quer ao nível de hardware, caminhado ao encontro das necessidades do seu utilizador [21].

De facto, os computadores são máquinas complexas, mas flexíveis e adaptáveis, assim melhorar a qualidade da interação Homem-computador é um dos campos de estudo da HCI. Existe um esforço para que a tecnologia se torne cada vez mais acessível a todos. Uma análise eficiente da interação entre o Homem e o computador pode ser de tal forma pertinente, que por exemplo, pode significar a diferença entre as pessoas utilizarem ou não determinada aplicação [13, 35].

Tal como o conceito indica, a HCI baseia-se na relação de três conceitos essenciais: o Homem como utilizador, o computador como máquina e a sua relação. De seguida serão discutidos cada um dos intervenientes assim como a sua relevância no estudo da HCI.

3.1.1 O Homem

O Homem, o utilizador final, é considerado a personagem central em qualquer sistema em que intervém, uma vez que este se apresenta como o elemento prioritário das suas interações. Os requisitos de qualquer sistema, devem, então, ser definidos em função deste, considerando as suas capacidades e limitações. Deve, ainda, ir ao encontro da forma como o Homem interpreta e se relaciona com o mundo ao seu redor, tal como a forma como este armazena e processa informações. Assim, a maneira como o Homem se relaciona com o computador pode ser entendida sob o ponto de vista de um modelo que engloba três sistemas:

- O Sistema perceptivo: este inclui os sentidos: a visão, o tacto, a audição, o paladar e o cheiro. Nesta interação específica com o computador, o paladar e o cheiro ficam um pouco de parte, uma vez que não assumem um papel preponderante.
- O sistema motor: que controla os movimentos oculares, cabeça, braços, membros. Destes elementos, os dedos têm o principal destaque, uma vez que lhes é atribuída a tarefa da digitação e manipulação do rato.

- sistema cognitivo: este sistema fornece o processamento necessário para conectar os dois anteriores. Deste sistema destaca-se a memória, aprendizagem e raciocínio. As informações recolhidas são armazenadas na memória, e posteriormente usada pelo raciocínio de forma a resolver os problemas. Situações familiares recorrentes permitem adquirir habilidades num determinado domínio à medida que as suas estruturas de informação se tornam mais definidas [21] .

Apesar do modelo apresentado incluir as principais características da interação do Homem com o computador, este pode ainda ser estendido e incluir mais duas componentes essenciais :

- Emoções: Os humanos são seres complexos que ultrapassam apenas as dimensões do cognitivo. As nossas ações são, em muito, influenciadas pelas nossas emoções. Por exemplo, emoções positivas incentivam-nos a resolver problemas mais complexos ou a ser mais criativos. Por outro lado emoções como a frustração ou o medo tendem a prejudicar o nosso desempenho. Desta forma, a maneira como nos relacionamos com o computador é influenciada pelas emoções que estamos a experienciar. Muitas das vezes o que nos leva a gostar de determinada aplicação é a sua interface, e em alguns casos isso deve-se às emoções positivas que estas, de algumas forma, nos está a transmitir, quer pela sua usabilidade, quer pelo seu lado estético.
- Perfil individual: No desenho ou conceção de um programa temos de assumir que todos os seres humanos são individuais. Assim deve-se considerar características como sexo, capacidades físicas, capacidades intelectuais e conhecimento informáticos . É útil, assim, ter em consideração o público alvo para um maior encontro entre o Homem e o computador.

3.1.2 O computador

Os computadores, tal como já referido, tratam-se de um objeto completamente popularizado nos dias de hoje, presente na maioria das nossas casas. Estes tratam-se de um dispositivo que pode ser visto de um ponto de vista mais tradicional como um conjunto de um ecrã, um teclado e um rato. No entanto os computadores hoje em dia entram nas nossas vidas com vários formatos, seja através dos *smartphones*, *tablets* ou "camuflados" sob a forma de um ecrã num frigorífico ou numa máquina de lavar roupa.

Para uma compreensão total de como os seres humanos interagem com o computador, é necessário compreender

esta interação como um todo. Nesta secção o foco será o computador mais tradicional e as suas interfaces mais comuns de interação: o rato e teclado uma vez que se trata do objeto de estudo da presente monografia. No entanto é necessário considerar que poderão existir muitas outras interfaces como são exemplo canetas *stylus*, impressoras, scanners, *webcam*, entre outros [21, 61].

3.1.2.1 Rato

O rato trata-se de um periférico de *input* que têm como principal função mover um apontador no pelo ecrã do computador. Este começou por possuir uma esfera na sua parte inferior que fazia girar dois discos no seu interior, permitindo mapear o meu movimento num eixo xy. Atualmente este tipo de funcionamento já se encontra ultrapassado, tendo sido substituídos por funcionamento ótico.

Geralmente os ratos apresentam quatro tipos de funções essenciais: clique simples, duplo clique, movimento e arrastar e largar. É com base nestas funções que assenta a interação com este dispositivo [21].

3.1.2.2 Teclado

Um teclado trata-se de um conjunto de teclas utilizado com o objetivo de introduzir dados no computador. Este possui números, letras e outros símbolos, onde cada tecla pode possuir um ou mais destes símbolos. Dependendo do sistema operativo, ou modelo, este pode apresentar algumas variações. Este dispositivo possui para além da função básica de digitação de texto, a possibilidade de aceder a uma vasta gama de atalhos [21].

3.1.3 A interação

Uma das principais teorias da interação Homem-computador foi apresentada por Norman em 1986. Este denominou esta teoria como "O ciclo de execução-avaliação", sendo composto por dois golfos: a da execução e a da avaliação [61]. O golfo da execução abrange o esforço cognitivo do utilizador para planejar a sua ação mediante as informações que foram recebidas do sistema. Por outro lado, o golfo da avaliação, trata-se do momento em que o utilizador avalia os resultados por comparação com o plano inicial e determina as suas novas ações. Entre estes dois golfos podem ser ultrapassados através de uma estratégia de sete passos:

Relativamente às sete fases referidas, estas apresentam-se como:

1. Estabelecer um objetivo;
2. Formular a intenção;
3. Especificar a sequência de ação;
4. Executar a ação;
5. Percecionar o estado do sistema;
6. Interpretar o estado do sistema;
7. Avaliar o estado do sistema relativamente às metas e objetivos;

Cada etapa apresentada trata-se de um exercício do utilizador. Este ciclo inicia-se com uma sequência de etapas 1 a 4 que culmina na ultrapassagem do primeiro golfo, o da execução. Assim, primeiramente, este formula um objetivo de alto nível (1). Estado o objetivo definido é necessário formular a intenção, ou seja, tomar a decisão de agir na busca do objetivo com base no atual estado do sistema(2). Posteriormente deve estabelecer quais as configurações do sistema que correspondem ao estado final desejado, especificando desta forma a sequência da ação. Isto exige um esforço mental de organização temporal das ações a posteriormente executar.(3) Neste ponto o utilizador encontra-se apto a executar a ação, ou conjuntos de ações de forma a atingir o objetivo (4). Após a execução da ação, sistema atualiza o seu estado e estamos perante o domínio do golfo da avaliação, onde devem ser executadas as etapas 5-7. Numa primeira fase, este irá percecionar quais as modificações ocorridas no sistema (5), posteriormente, o utilizador deverá interpretar essas modificações (6) de forma a avaliar se o estado do sistema está a ir ao encontro dos seus objetivos (7). Em caso positivo a interação foi bem sucedida. Caso contrário, o usuário deve formular uma nova meta e repetir o ciclo [21]. Este ciclo pode então ser resumidamente representado pela figura 3.1

3.2 HCI na avaliação de estados mentais

A forma como o Homem interage com o computador têm, ao longo do tempo, sido alvo de estudo de forma não só a melhorá-la, mas também com o intuito de extrair padrões de interação. Esses padrões depois de devidamente extraídos e processados podem ser úteis de forma a obter informações relevantes. Especificamente a forma como

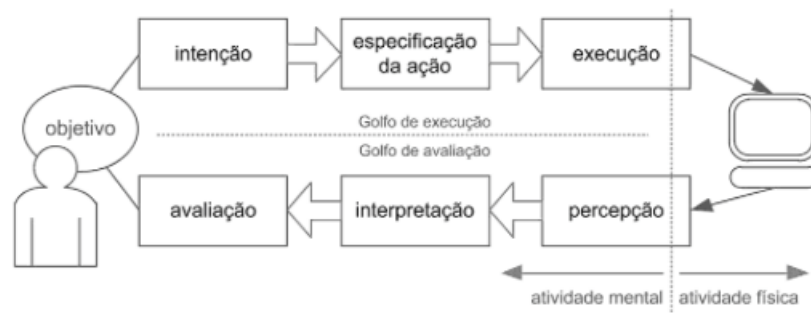


Figura 3.1: Etapas de atividades entre o golfo da execução e ação

o utilizador interage com o rato e o teclado têm sido estudada de forma a analisar a sua personalidade, deteção de emoções, estados afetivos e fadiga. Nestes estudos, são retiradas métricas como velocidade do rato, tempo de cliques, distâncias percorridas pelo rato, velocidades de digitação entre outras métricas. Os dados são posteriormente analisados de maneira a relacionar estes estados mentais com a interação com o rato e teclado. A utilização de padrões de HCI na deteção destes estados mentais, serão, então, abordadas mais especificamente nas próximas secções.

3.2.1 Emoções e estados afetivos

As emoções podem ser consideradas como todo um conjunto de reações biológicas, químicas e neurológicas que ajudam na sobrevivência dos seres vivos, originando comportamentos rápidos e eficazes. Desta forma, pode-se dizer que as emoções estão estritamente relacionadas com fatores comportamentais e afetam a tomada de decisão [18]. Assim, os computadores devem conseguir oferecer uma experiência adaptativa, capaz de lidar com as emoções e estados dos utilizadores. Em particular um estudo com base no movimento do rato foi apresentado, onde se verificou que movimentos mais rápidos por exemplo estavam relacionados com estados de entusiasmo.[70]. Também um estudo que apenas considerou interações através do teclado foi apresentado, que concluiu que 70% dos utilizadores diminuem a velocidade de escrita quando se encontram em estados emocionais negativos [38]. Mais tarde, em 2014, foi apresentado um outro estudo mas neste caso foram consideradas tanto as interações através do rato como as interações através do teclado. Neste último estudo foram recolhidos 42 métricas relativas ao teclado e 96 métricas relativas ao rato e relacionados os resultados com a escala de *Self-Assessment Manikin* (SAM) [56].

3.2.2 Personalidade

A personalidade pode ser vista como sendo um conjunto de padrões de comportamento consistente e processos interpessoais com origem num indivíduo. Assim, é do conhecimento comum a existência de diversas personalidades como pessoas introvertidas que preferem estar sozinhas, pessoas extrovertidos que se relacionam facilmente, a título de exemplo. A compreensão da personalidade pode ser útil, por exemplo, na conceção de interfaces personalizadas. Os estudos de HCI elaborados nesta área mostraram uma correlação entre a personalidade de cada pessoa e a forma como interagem com o rato e teclado. Nomeadamente foi possível detetar traços de personalidades como confiança, simpatia, excitação [36].

3.2.3 Fadiga Mental

A fadiga, o estado mental de interesse nesta dissertação, tal como já foi mencionado apresenta-se como um fenómeno complexo subjetivo e pessoal que pode ser refletida através de uma diminuição de desempenho ou diminuição da motivação. É também a responsável, em muitos casos por erros. O estudo da deteção fadiga mental, através do rato e teclado, foi iniciado como já referido em 2013. Neste estudo, como já referido, através de um conjunto de métricas inspiradas nas biométricas comportamentais foi possível detetar fadiga mental através da manipulação do rato e do teclado. Ainda foi provada uma relação direta entre a fadiga e uma diminuição da performance através do aumento do número de erros [50, 52, 51].

3.3 Conclusão

A interação do Homem com o computador trata-se de uma problemática atual e crescente nos dias de hoje, uma vez que o recurso a estas máquinas está cada vez mais enraizado nos nossos hábitos e dia-a-dia. O Homem tem um papel principal nesta interação, e o desenho dos computadores e das suas interfaces deve tornar a sua interação o mais natural e fluída possível.

Das interfaces mais tradicionais do computador, destacam-se o rato e o teclado. Estas interfaces já foram estudadas no passado de forma a reconhecer emoções, estados afetivos e fadiga mental. Nestes casos métricas como velocidade, tempo entre cliques, distâncias percorridas, tempo de pressão de teclas, velocidade de digitação foram

extraídos. Posteriormente a análise destas métricas permitiu reconhecer padrões.

Assim, conclui-se que através da recolha de padrões de HCI, quando devidamente extraídos e tratados, é possível a sua análise de forma a obter informações relevantes. Mais concretamente na área da fadiga mental, estudos prévios já conseguiram estabelecer uma relação entre esses padrões e estes estados mentais. É então possível afirmar que a deteção de fadiga através de padrões de HCI possível e viável.

Capítulo 4

Aprendizagem não supervisionada

A IA define-se como uma disciplina que tem como objetivo o estudo e construção de entidades artificiais com capacidades cognitivas semelhantes às dos humanos [46]. A inteligência é evidenciada nas mais variadas formas, seja pela compreensão de uma linguagem, raciocínio ou resolução de problemas. Mas sem dúvida que uma das formas mais explícitas de inteligência trata-se da capacidade de aprendizagem, ou seja o processo através do qual as competências, comportamentos, conhecimentos são alterados através da experiências, raciocínio e observação. Posto isto, facilmente se compreende a importância da aprendizagem no domínio da IA [54].

Geralmente aceita-se a existência de dois paradigmas centrais na aprendizagem : a aprendizagem supervisionada e a aprendizagem não supervisionada. A última abordagem distingue-se da primeira na medida em que o sistema aprende de forma independente e sem a intervenção de um “supervisor”. Este modo de aprendizagem consiste na utilização dos dados de entrada do sistema, não classificados, de forma a que a rede encontre as regularidades estatísticas entre eles [54].

Ao longo do presente capítulo serão distinguidos as duas formas de aprendizagem e aprofundado o conceito de aprendizagem não supervisionada, uma vez que se trata da forma de aprendizagem utilizada nesta dissertação.

4.1 Inteligência Artificial

A IA refere-se a um campo de estudo que nasceu em 1956, utilizado pela primeira vez por McCarthy [46]. Desde essa altura, e até aos dias de hoje diferentes correntes de pensamento têm tentado encontrar uma definição para este conceito. Estas correntes são geralmente divididas em quatro linhas de pensamento distintas:

1. Sistemas que pensam como seres humanos;
2. Sistemas que atuam como seres humanos;
3. Sistemas que pensam racionalmente;
4. Sistemas que atuam racionalmente.

De uma forma simplificada as linhas 1 e 3 baseiam-se no pensamento e raciocínio, enquanto que as linhas 2 e 4 baseiam-se no comportamento. Por outro lado, as correntes de pensamento 1 e 2 aceitam que se trata de uma busca com o objetivo da semelhança com o ser humano, enquanto que as restantes comparam com um ideal de inteligência denominado racionalidade. Apesar da distinção entre correntes, há um ponto comum entre elas, a reprodução de uma característica denominada por inteligência. Esta pode ser entendida como a habilidade para compreender e aprender novos conhecimentos, resolver problemas e tomar decisões. A inteligência manifesta-se nas mais diferentes formas, seja na compreensão de uma linguagem, no raciocínio ou na resolução de problemas [54, 72].

Introduzida a ideia de inteligência, compreende-se que ficou por esclarecer um conceito essencial para um sistema seja denominado de inteligente, ou possua inteligência: a sua capacidade de aprendizagem. A aprendizagem trata-se da capacidade de adaptação, de modificação das estruturas internas na busca da melhoria do comportamento, e consequentemente das suas respostas. Esta é sem dúvida uma das características mais importantes de um ser ou sistema inteligente [54].

Assim, aceita-se que o principal objetivo da IA, enquanto ciência, seja fazer com que as máquinas, ou sistemas, reproduzam comportamentos ou pensamentos que exigiriam inteligência quando executados por seres humanos.

4.1.1 Aprendizagem

Como já foi referido a capacidade de aprendizagem releva a inteligência de um sistema para aprender novos *inputs* e responder adequadamente aos mesmos. Assim, inspirados nas diferentes formas de aprendizagem evidenciadas nos seres humanos, convencionou-se que existem três principais formas de aprendizagem em sistemas inteligentes: aprendizagem não supervisionada, aprendizagem supervisionada e aprendizagem por reforço [46]. Estes conceitos serão mais aprofundados de seguida.

4.1.1.1 Aprendizagem por reforço

A aprendizagem por reforço apresenta-se como uma das vertentes da IA que permite que um agente computacional aprenda a partir da interação com o ambiente no qual se encontra inserido. Para tal recorre-se ao uso de recompensas. Este tipo de aprendizagem é especialmente útil quando se pretende obter um determinado comportamento, sem conhecimento da função que modela o mesmo [46, 14].

4.1.1.2 Aprendizagem supervisionada

Este trata-se do método de aprendizagem mais comum no treino das redes neuronais artificiais. Denomina-se de aprendizagem supervisionada porque a entrada e saída desejadas para a rede são fornecidas por um supervisor (também denominado professor). Desta forma o professor tem conhecimento sobre o ambiente indicando explicitamente o que é considerado um comportamento bom e mau para a rede [14].

O principal objetivo é ajustar os parâmetros da rede de forma a encontrar uma ligação entre os parâmetros fornecidos. As saídas calculadas são comparadas com as entradas solicitadas, é calculado o erro e este é comunicado à rede. O processo é repetido para cada entrada, ajustando-se os pesos das conexões, de forma a diminuir o erro. Quando o erro é inexistente podemos dispensar o uso do professor e deixar a rede neuronal lidar com o ambiente por si mesma.

A principal desvantagem deste tipo de aprendizagem assenta no facto de que na ausência de treinador a rede não consegue aprender novas estratégias para situações não abrangidas pelos exemplos de treinamento da rede [53].

4.1.1.3 Aprendizagem não supervisionada

Neste caso de aprendizagem não existe qualquer entidade ou professor que supervisione o processo de aprendizagem. Neste caso somente os padrões de entrada estão disponíveis para a rede. A partir do momento em que a rede estabelece uma harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente [53, 14].

Existem duas principais técnicas de aprendizagem não supervisionada: a associação e o *clustering* [68]. Sendo que esta monografia se foca na segunda técnica e nos seus algoritmos.

4.2 Clustering

Esta técnica é uma das principais abordagens ao problema da aprendizagem não supervisionada, o *clustering* ou análise de *clusters* têm como principal objetivo organizar os dados disponíveis em subgrupos homogêneos, também denominados de *clusters*, que partilham características semelhantes. Assim, procura-se encontrar a mínima variabilidade dentro dos *clusters* -homogeneidade- e a máxima distância -separação- entre os mesmos [68]. Deve-se igualmente ter em consideração que sendo uma técnica de aprendizagem não supervisionada, não há indicação sobre os membros dos grupos. Por estas razões a tarefa de análise de *clusters* pode tornar-se árdua e envolve uma análise detalhada dos grupos formados e das relações entre eles. De forma a ilustrar a natureza desta técnica considere-se



Figura 4.1: Exemplo de agrupamento [68]

a figura 4.1, onde podemos verificar a existência de 20 pontos. Tal como podemos averiguar, apenas neste exemplo pode-se encontrar 3 representações diferentes de agrupamentos dos objectos iniciais. A figura 4.2 apresenta uma divisão em 2 *clusters*, a figura 4.3 em 4 e por último a figura 4.4 apresenta 6 *clusters*. Não se pode afirmar que nenhuma esteja incorreta ou certa, são formas diferentes de agrupamento por isso considera-se que a forma mais

correta de agrupamento depende em muito da natureza dos dados, e do objetivo que se pretende atingir [3, 68].



Figura 4.2: Agrupamento com dois clusters [68] .



Figura 4.3: Agrupamento com quatro clusters [68].



Figura 4.4: Agrupamento com seis clusters [68] .

A aplicação desta técnica é vasta e contempla diferentes áreas, desde a biologia, meteorologia, psicologia, medicina, classificação de documentos, entre muitas outras. Posto isto, tendo em consideração todas as áreas de aplicação, a ambiguidade presente na representação dos *clusters*, e os diferentes tipos de dados, ao longo do tempo foram desenvolvidos diferentes medidas de similaridade, diversos algoritmos de *clustering* e distintas formas de avaliar a qualidade dos *clusters* [68]. De seguida serão abordados as medidas, algoritmos e métricas úteis à compreensão do caso de estudo.

4.2.1 Medidas de similaridade

Um dos tópicos que possui uma maior relevância na análise de *clusters* trata-se da definição de uma medida de distância entre os objetos, ou seja, uma medida de similaridade. São diversos os fatores que se deve ter em conta aquando da escolha de uma medida de similaridade, tal como a quantidade de dados, o tipo de dados e o domínio do problema.

As medidas de similaridade definem o grau de semelhança entre dois objetos e permitem posteriormente o agrupamento desses objetos com base no resultado obtido. Existe uma grande variedade de medidas de similaridade, no entanto nesta monografia serão abordadas as duas métricas mais usuais e mais citadas pela literatura: a Distância Euclidiana e a Distância de Manhattan [67, 68].

4.2.1.1 Distância Euclidiana

A distância Euclidiana é a distância mais utilizada e por isso a mais comum. Esta diz respeito à distância geométrica entre dois pontos baseada no teorema de Pitágoras. Quanto menor a distância entre esses pontos, maior a similaridade entre eles. Esta métrica é expressa pela equação 4.2.1

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4.2.1)$$

Em que a distância entre o ponto i e j é dado pela soma da raiz quadrada da diferença de cada coordenada. É de notar que a distância Euclidiana é fortemente afetada por variáveis de grandes dimensões ou com grandes dispersões, sendo que se apresenta imprecisa. Por este motivo é recomendada a normalização das variáveis de forma a diminuir o impacto das diferenças entre dimensões [67].

4.2.1.2 Distância Manhattan

A distância de Manhattan, contrariamente à distância à distância euclidiana, computa a distância necessária percorrida entre os dois pontos. Desta forma esta distância apresenta-se como a soma das diferenças entre i e j em cada

dimensão, representada pela seguinte equação:

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}| \quad (4.2.2)$$

Na figura 4.5 podemos verificar a diferença gráfica entre as duas métricas. A azul encontra-se assinalada a distância de Manhattan entre o ponto i e j, enquanto que a verde está assinalada a distância euclidiana entre os dois pontos já referidos [67].

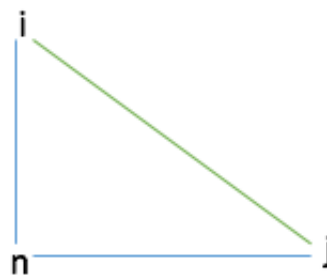


Figura 4.5: Distância de manhattan e distância euclidiana para o ponto i e j

4.2.2 Técnicas de clustering

Tal como já foi mencionado existem diversos algoritmos de *clustering* com diferentes propósitos. A escolha do algoritmo de *clustering* mais apropriado depende em muito do tipo de dados e do objetivo. Se a análise de *cluster* for usada como ferramenta de exploração é possível o teste de diferentes algoritmos de forma a descobrir o que levará a melhores resultados. Os algoritmos de *clustering* podem ser essencialmente divididos em 5 grupos: métodos de partição, métodos hierárquicos, métodos com base em densidade, métodos baseados em grelha e baseados em modelos. Os métodos hierárquicos e de partição serão o foco da presente monografia [31].

4.2.2.1 Métodos de partição

Os métodos por partição ou não hierárquicos baseiam-se no conceito que os objetos (dados) serão agrupados em k grupos, onde k é um número conhecido e definido previamente. É também convencionado que o número de k é menor

ou igual ao número de objetos existentes, sendo que desta forma cada *cluster* deve conter pelo menos um objeto.

Neste tipo de abordagem, depois de convencionado o número k chega-se ao número de partições iniciais, posteriormente o método aloca iterativamente os objetos até estarem agrupados numa "boa partição". Neste caso, uma boa partição é aquela em que os objetos dentro do mesmo *cluster* estão o mais próximo possível uns dos outros, enquanto que os objetos pertencentes a *clusters* diferentes estão o mais distante possível..

A função de minimização do erro mais frequentemente usada nos métodos particionais trata-se da soma dos quadrados dos erros, que é especialmente eficaz quando os *clusters* são compactos e bem separados. Assim, o erro quadrático para um agrupamento X , composto por um conjunto de objetos L é dado por:

$$e^2(X, L) = \sum_{j=1}^k \sum_{i=1}^{n_j} ||x_i^j - c_j||^2 \quad (4.2.3)$$

onde x_i^j é o i^{th} objeto pertencente ao j^{th} *cluster* e c_j trata-se do centróide do j^{th} do *cluster* [31].

De uma forma genérica existem dois principais métodos de partição que utilizam este erro como critério a minimizar. Estes distinguem na forma como obtém a melhor partição. Estes métodos, k-means e k-medoids, serão descritos de seguida.

k-means: O algoritmo k-means é de um dos algoritmos mais simples e por isso um dos mais utilizados. Este inicia-se pela escolha de k centros, onde k é um parâmetro definido pelo utilizador e como já referido reflete o número

de agrupamentos que se pretende obter como resultado. O algoritmo está, então, apresentado em 1.

Algoritmo 1: k-means

Entrada: O número de grupos k e os N objetos

Saída: k grupos

1 início

2 Escolher de forma aleatória k objetos do conjunto inicial, como os centros iniciais dos grupos;

3 repita

4 Atribuir cada elemento ao grupo ao qual o elemento é mais similar (mais próximo), de acordo com o valor médio dos elementos no grupo;

5 Atualizar as médias dos grupos, calculando o valor médio dos elementos para cada grupo;

6 *até que não haja mudanças de elementos de um grupo para outro, ou minimizar o erro calculado em 4.2.3;*

7 fim

Devido à forma como se inicializa este algoritmo, o *k-means* pode não convergir para a melhor partição, uma vez que a escolha inicial dos k objetos que vão constituir os centros iniciais pode influenciar o resultado. Desta forma os grupos obtidos apenas representam um local ótimo da função de avaliação para aquela inicialização. Uma forma de contornar este problema é repetir o algoritmo para distintos centros iniciais. Também são dadas como soluções uma escolha criteriosa dos centros, ou a utilização de outros algoritmos de *clustering* como os métodos hierárquicos [68].

Relativamente à complexidade, este é um algoritmo na ordem de tempo de $O(l k n)$ e uma complexidade de espaço na ordem $O(k)$, onde n refere-se ao número de dados e l o número de iterações necessárias até que não haja mudanças. Isto significa que a nível de espaço, este é assim considerado um algoritmo modesto, da mesma forma também não é computacionalmente exigente ao nível de tempo, sendo o que apresenta um comportamento linear relativamente ao número de objetos (visto que k e l são geralmente fixados à priori) [31].

Algumas das características gerais deste agrupamento são a sua sensibilidade ao ruído, uma vez que trabalha com médias e ainda sua tendência para formar grupos esféricos e de tamanho semelhante [68].

k-medoids Este método contrariamente ao k-means utiliza um ponto central para representar o centro do *cluster*, sendo que este ponto se denomina de medoide. Assim, o objetivo à semelhança do k-means é encontrar um medoide representativo para cada k , onde cada objeto é atribuído ao medoide mais semelhante. O algoritmo relativo a este método encontra-se descrito em 2.

Algoritmo 2: k-medoid

Entrada: O número de grupos k e os N objetos

Saída: k grupos

1 início

2 Escolher de forma aleatória k objetos do conjunto inicial, como os centros iniciais dos grupos;

3 repita

4 Atribuir cada elemento ao grupo ao qual o elemento é mais similar (mais próximo), de acordo com medoide do grupo;

5 Aleatoriamente, selecione um elemento que não esteja como medóide, r ;

6 Calcule o custo total de trocar o medóide atual pelo elemento r ;

7 Se o custo for menor então troque o medóide atual por r

8 **até** que não haja mudanças de elementos de um grupo para outro.;

9 fim

Algumas das principais características deste agrupamento são, à semelhança do *k-means*, a sua tendência para encontrar grupos esféricos. No entanto possui um processamento mais custoso que o anterior método e por isso não é aplicável a grandes bases de dados. Ainda assim, é mais robusto que o k-means a lidar com *outliers*. Uma das formas alternativas para aplicação do algoritmo k-medoids a grandes bases de dados é o algoritmo CLARA uma vez que utiliza apenas uma parte dos objetos como uma amostra representativa da população e escolhe nesses dados os medóides. No caso da amostra ser selecionada de forma aleatória, esta deverá representar bem o conjunto de dados originais, e desta forma garante bons resultados.

Fuzzy c-means A grande parte das abordagens de *clustering* definem que cada objeto pertence unicamente a um grupo. No métodos de *fuzzy*, define-se que cada objeto pode ter um grau de pertinência a cada *cluster*. Sendo assim, os objetos não pertencem rigidamente a um agrupamento. Este grau de pertinência varia entre 0 e 1, dado que quando perto de 1 o objeto é similar ao seu *agrupamento* e 0 não o é.

Dos métodos de *fuzzy*, o mais bem estudado e mais aplicado trata-se do fuzzy-c-means, uma extensão do método k-means para o domínio de *fuzzy*. [66] Neste caso, a função a minimizar trata-se de :

$$J(X) = \sum_{i=1}^n \sum_{k=1}^K u_{ik}^m \cdot ||x_i - c_k||^2 \quad (4.2.4)$$

onde X representa o índice de desempenho do agrupamento, n o número total de objetos do conjunto, x_i é o objeto i desse conjunto, c_k representa o centro do agrupamento k e o u_{ik} trata-se do valor de pertinência do objeto i ao agrupamento k . Assim, esta função permite obter o valor das somas de cada objeto ao centro do *cluster*. No entanto, de forma a completar a equação acima é necessário encontrar o grau de pertinência que é atualizado a cada iteração. A atualização de cada grau de pertinência segue a seguinte equação:

$$u_{ik}^{t+1} = \left[\sum_{j=1}^K \left(\frac{||x_i - c_k^t||^2}{||x_i - c_j^t||^2} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (4.2.5)$$

Enquanto que cada centro de cada agrupamento é dado por:

$$c = \frac{\sum_{k=1}^n u_{ij}^m \cdot x_k}{\sum_{k=1}^n u_{ij}^m} \quad (4.2.6)$$

Algoritmo 3: Fuzzy-c-means

Entrada: N objetos, k grupos e expoente de fuzzilização m, matriz de partição aleatória inicial U e ξ critério de paragem

Saída: k grupos, matriz de pertinência U

1 início

2 repita

3 Calcular o centro dos k agrupamentos pela equação 4.2.6

4 Atualizar a matriz de similaridade U^{t+1} onde t é a iteração atual, segundo a equação 4.2.5

5 **até** Calcular $\varepsilon = ||U^{t+1} - U^t||$, Se $\varepsilon < \xi$ final do algoritmo;

6 fim

4.2.2.2 Métodos hierárquicos

Os métodos hierárquicos criam uma hierarquia entre os conjuntos de objetos. Existem duas formas distintas de agrupamento: a aglomerativa ou *bottom-up* e a divisiva ou *top-down*. A primeira distingue-se da segunda na medida em que inicia a decomposição hierárquica a partir de um conjunto de elementos isolados, enquanto que a segunda começa com um grande conjunto e vai dividindo-os em partes até obter elementos isolados. Neste tipo de método não é necessária a definição do número de grupos [68] [3].

Nos métodos hierárquicos, os grupos são frequentemente representados por diagramas denominado dendograma. Na figura 4.6 pode-se verificar um exemplo do mesmo. Neste exemplo de diagrama cada ramo representa um objeto, e quanto mais se sobe, menos divididos os objetos estão. A tracejado azul, encontra-se um exemplo de uma distância de corte. Esta distância de corte pode ser especificada conforme o objetivo final. Neste exemplo em concreto, a distancia de corte origina três *clusters* [3].

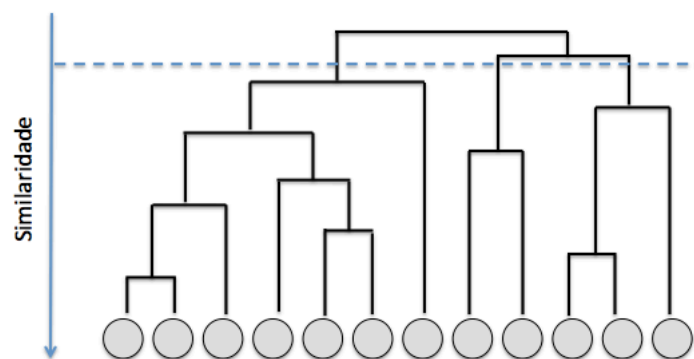


Figura 4.6: Exemplo de dendograma

Métodos aglomerativos Estes são dos métodos mais comuns dentro dos métodos hierárquicos. Tal como já referido, nesta abordagem cada elemento inicia no seu agrupamento, e a cada passo, os grupos vão-se ligando de acordo com a sua similaridade, obtendo por fim, um grande grupo com todos os elementos.

Uma das características destes métodos, trata-se, de que uma vez dado um passo de divisão ou junção, este já não

é reversível, sendo o principal responsável pelo seu baixo custo. No entanto, o cálculo das matrizes de similaridade, dependendo do tamanho, podem requer muita memória e tempo de processamento.

De forma geral, o procedimento geral para os algoritmos *aglomerativos* pode ser descrito pelo algoritmo 4.

Algoritmo 4: Método hierárquico aglomerativo

Entrada: n objetos

Saída: Conjunto de grupos

1 início

2 Iniciar o algoritmo com n grupos, sendo que cada elemento representa um grupo e uma matriz de similaridade $D_{n \times n}$

3 repita

4 Localizar a menor distância, ou seja a maior similaridade entre o objeto x_1 e x_2 , $D_{x_1 x_2}$;

5 atualizar a matriz de similaridade D , retirando os objetos x_1 e x_2 ;

6 atualizar a matriz adicionando as novas distancias do grupo encontrado (x_1, x_2) ;

7 **até** $n-1$, quando todos os objetos pertencem ao mesmo grupo;

8 fim

Existe uma variedade de algoritmos aglomerativos, sendo os mais utilizados os *single-link* e *complete-link*. Estes distinguem-se na forma como definem as distâncias entre grupos. No primeiro caso é definida como a menor distância entre os objetos, enquanto que no *complete-link* é calculada a maior distância. [68] [31]

Métodos divisivos Estes tratam-se dos métodos menos utilizados, uma vez que são pouco eficientes e exigem um esforço computacional superior aos métodos aglomerativos. Contrariamente aos anteriores, estes iniciam-se com todos os objetos juntos no mesmo grupo, e vão gradualmente sendo divididos até que exista um objeto por agrupamento. Este algoritmo encontra-se descrito em 5 .

Como se pode verificar a exigência da divisão de todas as possíveis combinações dos dados em dois *clusters* pode ser computacionalmente muito exigente, sobretudo quando há um grande número de dados. Nesses casos pode ser quase impraticável.

Algoritmo 5: Método hierárquico divisivo**Entrada:** N objetos**Saída:** Conjunto de grupos**1 início****2** | Inicia-se com um único grupo contendo todos os objetos**3 repita****4** | Computa-se a matriz de similaridade D entre todos os possíveis pares de agrupamentos**5** | Ocorre a formação de um novo agrupamento pela divisão dos objetos com menor grau de similaridade**6** | **até** *os passos anteriores até que no final se obtenha um objeto por cluster;***7 fim****4.2.3 Avaliação dos clusters**

A avaliação dos *clusters* têm como objetivo final aferir sobre a qualidade dos agrupamentos encontrados. Essencialmente, esta avaliação pode ser feita recorrendo a métricas internas ou métricas externas. Enquanto que as métricas internas medem a qualidade com base nos agrupamentos encontrados, não utilizando para isso informações externas, as métricas externas utilizam os classes pré-definidas para aferir se estes encontram-se bem ou mal agrupados.

4.2.3.1 Métricas internas

- Índice de Silhueta: Cada *cluster* pode ser representado como uma silhueta , este índice fornece informações sobre a coesão e separação dos agrupamentos. A silhueta de um *cluster* pode ser dada como:

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

onde a_i representa a dissimilaridade média entre o objeto i em relação aos restantes objetos do *cluster* A ao qual i pertence, b_i trata-se da dissimilaridade média do objeto i em relação aos objetos do *cluster* que lhe está mais próximo e é denominado de B. Um valor de s_i próximo de 1 indica que os dados estão bem agrupados e separados, se próximo de 0 indica que os agrupamentos encontrados podem não ser estáveis. Por último, quando perto de -1 indica que os dados estão mal agrupados. A largura média geral da silhueta para todo o conjunto de dados é a média de s_i para todos os pontos do conjunto de dados.

- Índice de Davies-Bouldin: Este índice tenta minimizar a distância média entre cada *cluster* e o mais semelhante

a ele. É dado por:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} = \frac{\text{diam}(c_i) + \text{diam}(c_j)}{d(c_i, c_j)}$$

Pequenos valores de DB correspondem a grupos mais compactos.

- Índice de Dunn: Este índice, tenta encontrar os *clusters* mais compactos e bem separados. Um índice de Dunn mais elevado corresponde a *clusters* melhor separados. Este é dado por:

$$ID = \frac{d(C_i C_j)}{\text{diam}(C_h)}$$

onde $d(C_i C_j)$ representa a distância entre C_i e C_j representa e C_h trata-se da classe com o maior diâmetro.

4.2.3.2 Métricas externas

- Precisão e Revocação: A precisão e a revocação são duas métricas amplamente usadas na avaliação externa da qualidade dos resultados nas mais variadas áreas. Considere-se para tal a matriz de confusão expressa na tabela 4.1.

Tabela 4.1: Exemplo de matriz de confusão

	Saudável	Doente
Diagnóstico saudável	Verdadeiro Positivo	Falso Positivo
Diagnóstico doente	Falso Negativo	Verdadeiro Negativo

Nesta matriz é considerado o exemplo do diagnóstico de pacientes contra a classificação real de saudável ou doente. Assim, um verdadeiro positivo é aquele diagnosticado como saudável e que realmente está saudável. Um verdadeiro negativo são aqueles que estão doentes e o diagnóstico coincidiu. Por outro lado, um falso negativo é aquele cujo o diagnóstico era doente mas afinal revelou-se estar saudável, por último um falso positivo é aquele cujo o diagnóstico era saudável mas afinal deu-se como doente.

Com base nestas aferições é possível calcular a precisão e a revocação. [9] Sendo que a primeira é dada como :

$$Preciso = \frac{VP}{VP + FP}$$

Enquanto que a revocação é dada como [9]:

$$Revocao = \frac{VP}{VP + FN}$$

- F1: Esta métrica é dada pela média harmónica ponderada da precisão e revocação e é dada como [9]:

$$F_1 = \frac{2 * P * R}{R + P}$$

4.3 Conclusão

A técnica de *clustering* trata-se de uma abordagem não supervisionada onde os algoritmos são agrupados de forma a eles próprios criarem as suas próprias classes produzindo vários subgrupos, segundo a especificação do utilizador. A grande diferença relativamente à aprendizagem não supervisionada prende-se no facto de que estes não necessitam de classes já previamente definidas.

Para aplicação desta técnica existe um largo espectro de algoritmos sendo que os abordados nesta monografia se prendem aos algoritmos de partição e hierárquicos. Enquanto que no primeiro os utilizadores devem referenciar um parâmetro k, que indica o numero de subgrupos a procurar, no caso do segundo método não existe a necessidade deste parâmetro e os algoritmos seguem uma abordagem *bottom-up* e *top-down*.

A tarefa de *Clustering* trata-se de uma operação que requer uma exploração de dados muito exigente, sendo que por isso trata-se de uma tarefa complexa e temporalmente dispendiosa.

Capítulo 5

Caso de Estudo

Como já foi apresentado, a presente dissertação tem como objetivo um estudo da aplicação de algoritmos não supervisionados com o intuito de detetar casos de fadiga mental. Para a concretização deste objetivo, foi proposto a utilização de padrões de HCI, recolhidos através da sensorização de rato e teclado. Assim, um conjunto de dados, do domínio em questão, foi cedido pela *startup Performetric* para ser o ponto de partida deste estudo.

A abordagem utilizada passou por um estudo detalhado dos dados cedidos, seguidos da aplicação de aprendizagem não supervisionada para resolução do problema proposto. Os resultados foram posteriormente comparados com metodologias já estudadas e implementadas.

Neste capítulo é apresentada a metodologia e incide ainda sobre toda a preparação dos dados, assim como o estudo aprofundado do *dataset*.

5.1 Metodologia

A metodologia adotada para a resolução do problema passou por várias fases sequências que culminaram na deteção de fadiga mental. Para isso foi utilizada uma abordagem iterativa baseado no processo de aprendizagem, adaptado à problemática em questão. Desta forma, a metodologia divide-se em 4 fases:

- **Preparação do dados:** Esta tratou-se da primeira fase do processo, e possuiu um impacto bastante significativo no resultado final. Nesta fase foram selecionados os dados com relevância para o problema, sobre os quais os algoritmos de aprendizagem não supervisionada incidiram. Foi igualmente durante esta fase que os dados sofreram uma limpeza e a eliminação de *outliers*, uma vez que é normal o aparecimento de dados redundantes, incoerentes e observações atípicas.

Desta forma compreende-se a necessidade de uma abordagem bastante cautelosa durante esta fase, uma vez que a qualidade dos dados resultantes vai determinar a eficiência dos algoritmos aplicados.
- **Aprendizagem não supervisionada:** Nesta etapa foram aplicados os algoritmos não supervisionados de forma a descobrir padrões existentes nos dados. Os algoritmos escolhidos nesta fase possuem igualmente uma enorme importância, uma vez que estes devem adequar-se ao tipo e à estrutura que os dados apresentam.
- **Interpretação e avaliação:** Neste ponto os resultados dos diferentes algoritmos foram interpretados e comparados entre si de forma a tirar conclusões sobre os agrupamentos encontrados. Os resultados foram igualmente comparados com avaliações subjetivas dadas pelos utilizadores sobre o seu nível de fadiga.
- **Aplicação do modelo a novos casos:** Após uma avaliação dos agrupamentos encontrados, foram selecionados aqueles que apresentaram os melhores resultados com o objetivo que estender o modelo na deteção de novos casos de fadiga.

5.2 Preparação dos dados

O ponto de partida do desenvolvimento da parte experimental da presente dissertação tratou-se, como já mencionado, de um *dataset* fornecido pela *startup* Performetric. O *dataset* cedido resultou da monitorização de utilizadores familiarizados com o uso corrente do computador. Esta monitorização, foi totalmente não invasiva e traduziu-se na recolha de métrica relacionadas com a interação dos utilizadores com o rato e teclado. Ao longo da interação dos utilizadores com o computador, a aplicação correu em *background* sem qualquer perturbação das tarefas que estavam a ser executadas, sendo que foram recolhidas oito métricas de 5 em 5 minutos. As métricas foram recolhidas, durante o período determinado, com base em duas componentes estatísticas: a média e a variância. O *dataset* apresentava, então, as seguintes métricas:

- *Key Down Time (KDT)* : Considera a etiqueta temporal para a pressão de uma tecla.
- *Time between keys (TBK)*: Etiqueta temporal decorrido entre pressionar duas teclas.
- *Mouse velocity (MV)*: Velocidade do rato, medida em pixels por milissegundos.
- *Mouse acceleration (MA)*: Aceleração do rato, ou seja velocidade (pixels/milissegundos) pelo tempo (milissegundos).
- *Time between clicks (TBC)* : Etiqueta temporal decorrido entre dois clicks do rato.
- *Double click duration (DCD)* : Etiqueta temporal decorrido entre duplos clicks do rato.
- *Average excess of distance (AED)*: É dado pela divisão entre distância real percorrida pelo rato e distância em linha reta entre os dois clicks
- *Average distance of the mouse to the straight line (ADMSL)*: Esta medida oferece a distância média entre as coordenadas do caminho real que o rato percorre e o ponto mais próximo desse caminho em linha reta.
- *Distance of the mouse to the straight line (DMSL)*: É calculado da mesma forma que a variável anterior, mas neste caso em vez de se calcular a média da distância entre o caminho real e o caminho em linha reta, apresenta-se o somatório das distâncias.

Para cada uma das métricas supramencionadas, o *dataset* inclui, ainda, uma autoavaliação aferida pelos próprios utilizadores para cada um dos vetores de dados recolhidos. Para a concretização desta autoavaliação, foi pedido aos utilizadores que se classificassem numa escala de sete níveis sobre o seu estado de fadiga mental, seguindo desta forma os níveis propostos pela escala USAFSAM.

Além do resultado da autoavaliação, o *dataset*, ainda incluía a tarefa que o utilizador estaria a desempenhar no momento da recolha das métricas. No momento da inscrição do utilizador no sistema de monitorização, foram igualmente recolhidas informações pessoais sobre os mesmos. Estas informações incluíam o nome, a data de nascimento, o género e a ocupação. [50]

5.2.1 Processamento dos dados

Numa primeira fase, após a obtenção do *dataset*, foi necessário proceder a uma limpeza do mesmo, uma vez que foi possível detetar algumas incoerências nos dados entre as quais: dados duplicados, entradas classificadas com níveis de fadiga incoerentes (como por exemplos níveis negativos), e ainda métricas com valores nulos.

Concluída a limpeza de valores incoerentes, partiu-se para a deteção de *outliers*. Apesar de estar implementado pelo sistema de recolha e transformação dos dados da *Performetric* um primeiro tratamento de valores anómalos, foi ainda assim verificada a existência de valores aberrantes nas métricas em estudo, tal como pode ser verificado na figura 5.1, onde a título de exemplo se apresenta um gráfico de caixa onde é possível observar a existência de valores de ADMSL atípicos.

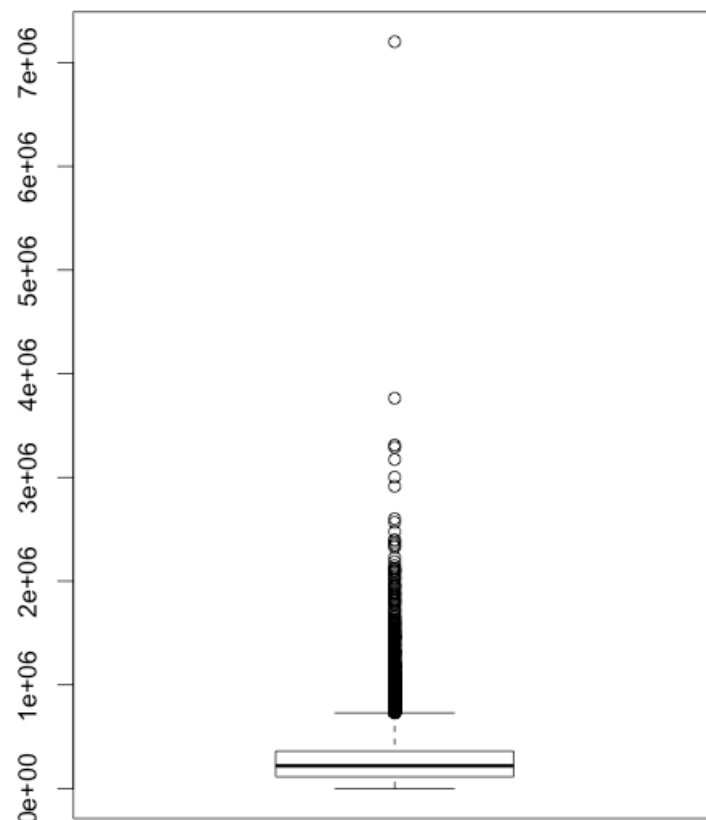


Figura 5.1: Diagrama de caixa para a métrica ADMSL

Posto isto, foi necessário definir uma abordagem para a deteção dos *outliers*. Uma vez que se tratam de dados

provenientes de diferentes utilizadores, em diferentes níveis de cansaço, uma simples abordagem de eliminação dos valores extremos aberrantes pode não ser suficiente. Assim sendo, optou-se por uma abordagem diferente. Uma vez que cada métrica era composta por duas componentes: média e variância, optou-se pelo cálculo do coeficiente de variação de Pearson expresso pela equação:

$$c_v = \frac{S}{\bar{x}} \quad (5.2.1)$$

onde S se trata do desvio padrão e \bar{x} trata-se da média [7]. Assim, foi possível obter uma medida de dispersão relativa que permitiu aferir se durante o intervalo em que foram extraídos os dados, a média obtida é o resultado de valores homogêneos, ou se pelo contrário existiram valores muito heterogêneos. Esta informação pode ser muito útil uma vez que altos valores de variância podem indicar instabilidade das métricas, e poderão afetar o estudo em causa. Após o cálculo desta medida para cada uma das métricas, os dados foram organizados por utilizadores e para cada grupo formado foi aplicado o algoritmo *pcout*, disponibilizado na biblioteca do *mvoutlier*¹. Este algoritmo de remoção de *outliers* é especialmente útil na tarefa de lidar com *datasets* com multi variáveis, uma vez que as técnicas mais usadas como a identificação com recurso a gráficos caixa apenas considera uma dimensão das variáveis de cada vez.

5.3 Pré análise dos dados

O *dataset* final é então proveniente de 18 utilizadores distintos, 10 homens e 8 mulheres, com idades compreendidas entre os 23 e os 50 anos, composto por 77 entradas, 11 de cada nível. O resumo das métricas em estudo está então apresentado na tabela 5.1

Tabela 5.1: Síntese dos dados

Variável	Média	Mediana	Valor máximo	Valor mínimo
KDT	90.69	96.10	242.79	0.00
MA	0.64158	0.64677	1.61946	0.04477
MV	0.61367	0.58680	1.66119	0.09049
TBC	2710	1775	20412	0.00
DDC	160.46	101.07	984.85	0.00
DMSL	25416171	20474550	128572013	0.00
AED	2.048	1.845	10.677	0.00
ADMSL	223366	204408	805350	0.00

¹<https://cran.r-project.org/web/packages/mvoutlier/mvoutlier.pdf>

Com um intuito de fazer um estudo da influência da fadiga nas métricas anteriormente apresentadas, optou-se por dividir inicialmente o *dataset* em dois grupos com base nas respostas da autoavaliação do estado de fadiga aferidos pelos utilizadores. Os grupos resultantes foram denominados de "não fatigado" e "fatigado" e equivalem respetivamente aos níveis 1 a 3 e níveis 4 a 7 da escala USAFSAM. O principal objetivo desta abordagem foi compreender se todas as métricas em estudo eram realmente influenciadas pela fadiga. Uma vez que não se está a lidar com nenhuma distribuição particular dos dados, foi necessário recorrer a um teste não paramétrico.

Sendo assim, o teste adotado tratou-se do teste de Mann-Whitney. Este teste não paramétrico é aplicado a amostras independentes. Neste caso, quis-se provar a independência entre as amostras provenientes de estados mentais de fadiga e estados mentais normais. O teste de Mann-Whitney assenta no teste de hipótese, onde posteriormente é tomada uma decisão de aceitar ou rejeitar a hipótese. Desta forma, a hipótese nula assumida neste teste é a seguinte: H_0 as medianas das duas distribuições são iguais, sendo que por cada uma das amostras em comparação é retornado um p-value, que quanto menor, menor a probabilidade de ocorrer a hipótese nula. Nos testes efetuados assumiu-se um nível de significância α , de 0.05, sendo que quando o $p\text{-value} < \alpha$ considerou-se a existência de diferenças entre os grupos e a H_0 foi rejeitada. Desta forma recorreu-se a este teste para inferir a existência de uma diferença significativa entre os estados mentais normais e estados de fadiga, e assim verificar se a variável em questão é realmente um bom indicador de fadiga.

Na tabela 5.2 pode-se observar o resultado obtido da aplicação do teste de Mann-Whitney para cada uma das componentes em estudo. Pode-se verificar que as componentes de MA e MV tratam-se das métricas com mais diferença significativa entre os dois grupos, seguindo-se da ADMSL, KDT, DDC e DMSL que também tiveram o mesmo comportamento, uma vez que apresentaram um $p\text{-value} < 0.05$. Relativamente às métricas do TBC e DMSL o teste revelou que a H_0 foi aceite, e que esta métrica não têm diferença significativa entre o *dataset* de estados mentais não fatigados, e estados mentais de fadiga.

Uma vez verificada a existência de uma relação entre os dois estados em consideração e as métricas em estudo, realizou-se um estudo equivalente para os sete níveis aferidos pelos utilizadores. Para o efeito recorreu-se ao teste de Kruskal-Wallis, uma extensão do teste de Mann-Whitney que permite a comparação de duas ou mais distribuições. À semelhança do teste de Mann-Whitney a sua função é testar a hipótese nula de que as populações em estudo possuem a mesma distribuição, contra a hipótese que pelo menos duas distribuições têm distribuições diferentes. Desta forma,

Tabela 5.2: Resultado da análise de Mann-Whitney

Variável	p-value	H_0
KDT	0.0176	Rejeitada
MA	0.0016	Rejeitada
MV	0.0004	Rejeitada
TBC	0.2571	Aceite
DDC	0.0052	Rejeitada
DMSL	0.0691	Aceite
AED	0.0007	Rejeitada
ADMSL	0.0152	Rejeitada

o *dataset* foi reorganizado em 7 grupos de 11 instâncias de cada nível (de 1 a 7), em que foi posto à prova a seguinte hipótese nula H_0 : *as medianas das distribuições são iguais*. De forma semelhante ao teste anterior é retornado um *p-value*, que quanto menor, menor a probabilidade da hipótese nula se verificar. Assumiu-se igualmente um α de 0.05, sendo que quando se verificou a premissa $p\text{-value} < \alpha$ a hipótese foi rejeitada.

Tabela 5.3: Resultado da análise de Kruskal-Wallis

Variável	p-value	H_0
KDT (média)	0.4118	Aceite
MA (média)	0.0461	Rejeitada
MV (média)	0.0144	Rejeitada
TBC (média)	0.1208	Aceite
DDC (média)	0.1051	Aceite
DMSL (média)	0.0563	Aceite
AED (média)	0.0029	Rejeitada
ADMSL (média)	0.0151	Rejeitada

Como se pode analisar na tabela 5.3 a hipótese nula foi rejeitada no caso da MA, MV, AED e ADMSL. Isto significa que estas métricas apresentam as diferenças mais significativas entre os grupos, enquanto que nas restantes não foram encontradas diferenças segundo a hipótese formulada.

Considerando-se os resultados obtidos, foram analisados detalhadamente cada um das distribuições em função das métricas para o qual a Hipótese Nula foi rejeitada. Optou-se então, por não considerar na análise de estados de não fadiga mental vs estados de fadiga mental as métricas TBC e DMSL, uma vez que como já estudado não teriam influência na distinção dos grupos. Relativamente à análise segundo a escala USAFSAM, optou-se por não se considerar

as variáveis KDT, TBC, DDC e DMSL, pelas mesmas razões. De seguida apresenta-se a análise individual de cada uma das variáveis.

5.3.1 Key Down Time

Esta métrica, tal como já foi referido, mede o tempo de pressão de uma tecla. É expectável que quanto mais fatigados os utilizadores se apresentem, maior seja o KDT obtido. Isto deve-se ao facto que um movimento mais lento pode estar associado a estados de fadiga, resultando, desta forma, num maior tempo de pressão das teclas.

5.3.1.1 Ausência de fadiga vs estado de fadiga mental

A figura 5.2 e tabela 5.4 apresentam o resultado da análise dos dois estados mentais: não fatigado e fatigados. Assim, tal como esperado, é notório um aumento da média de 78.78 milissegundos, no grupo "não fatigado", para 98.99 milissegundos para o grupo de "fatigado". Este aumento também foi verificado na mediana, de 83.01 para 102.73 e no valor máximo extremo.

Os resultados experimentais verificam, assim, as suposições iniciais, verificando que um maior tempo de pressão das teclas está associado a um estado de fadiga mental.

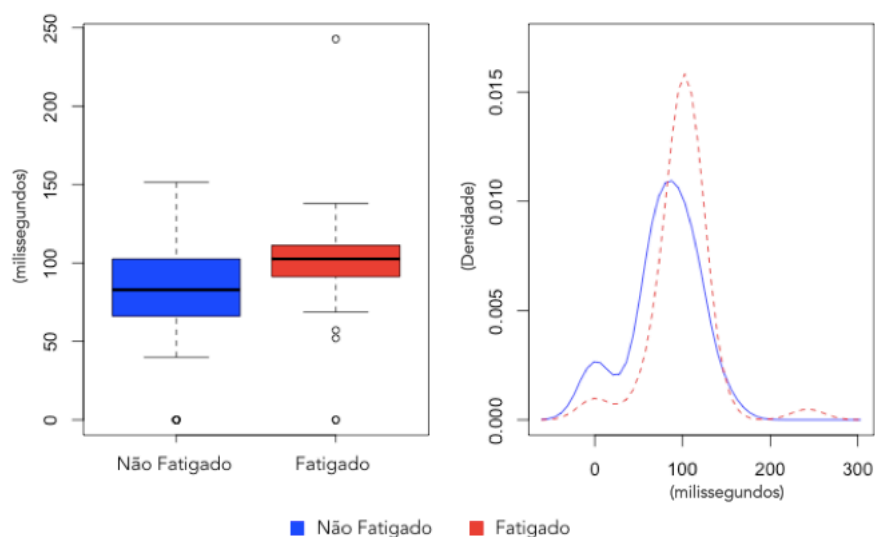


Figura 5.2: Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao KDT

Tabela 5.4: Valores médios dos dados para o KDT

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	78.78	83.04	151.64	0.00
Fatigado	98.99	102.73	242.79	0.00

5.3.2 Mouse acceleration

Relativamente à aceleração do rato é expectável que esta métrica possua menores valores nos estados de fadiga mental, uma vez que este movimento se encontra relacionado com a reação, memória e precisão associados, por sua vez, à diminuição de performance caraterísticos de estados de cansaço.

5.3.2.1 Ausência de fadiga vs estado de fadiga mental

Confrontados os grupos "fatigado" e "não fatigado", conclui-se que contrariamente ao que seria esperado o gráfico 5.3 e tabela 5.5 evidencia um claro aumento da MA em estados de fadiga mental. Este aumento é observado na média, que sofre um acréscimo de $0.52450 \text{ pixels/milissegundo}^2$ para $0.72905 \text{ pixels/milissegundo}^2$, assim como na mediana que sofre um incremento muito semelhante. Tal comportamento também é evidenciado pelo aumento do ponto de extremo máximo.

Assim, conclui-se que neste caso o comportamento de uma pessoa fatigada, contrariamente ao previsto não sofre uma diminuição na aceleração do rato, mas sim um aumento.

Tabela 5.5: Valores médios dos dados para o MA

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	0.52450	0.52341	1.03432	0.04477
Fatigado	0.72905	0.72148	1.61946	0.06929

5.3.2.2 Escala USAFSAM

A tabela 5.6 e figura 5.4 evidenciam a análise dos 7 níveis de fadiga mental para a métrica em estudo. Verifica-se que não existe uma relação linear entre a aceleração do rato e os níveis de fadiga, sendo que o nível que apresenta os valores médios e medianos mais baixos trata-se do nível 3, enquanto que o nível que apresenta os valores médios e medianos mais elevados trata-se do nível 5.

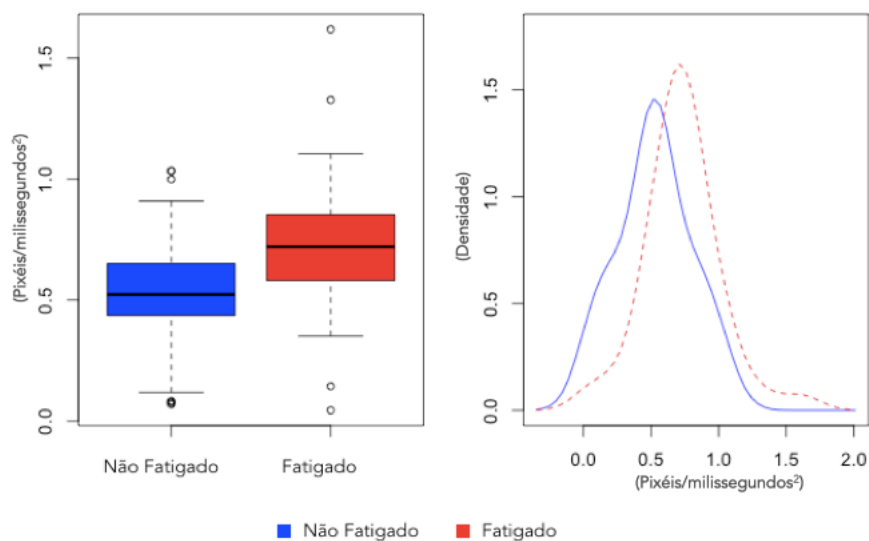


Figura 5.3: Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao MA

Tabela 5.6: Análise dos grupos USAFSAM para a métrica MA

Nível de fadiga mental	Média	Mediana	Máximo	Mínimo
1	0.52311	0.48847	1.03432	0.08145
2	0.59930	0.56316	1.00006	0.07477
3	0.45248	0.48649	0.87741	0.06929
4	0.60916	0.68911	1.05801	0.04477
5	0.8165	0.7510	1.6195	0.4230
6	0.7092	0.6468	1.0687	0.5105
7	0.7814	0.7305	1.3281	0.5954

5.3.3 Mouse velocity

Uma vez que a velocidade e aceleração tratam-se de dimensões que estão relacionadas, espera-se um comportamento semelhante. Tal como na aceleração do rato, seria previsível a obtenção de uma velocidade inferior nos casos de fadiga mental.

5.3.3.1 Ausência de fadiga vs estado de fadiga mental

Tal como ocorreu com a aceleração do rato, na velocidade do rato é notório um aumento da média, mediana, e pontos extremos. O que indica, que contrariamente ao suposto inicialmente, a velocidade do rato aumenta com o

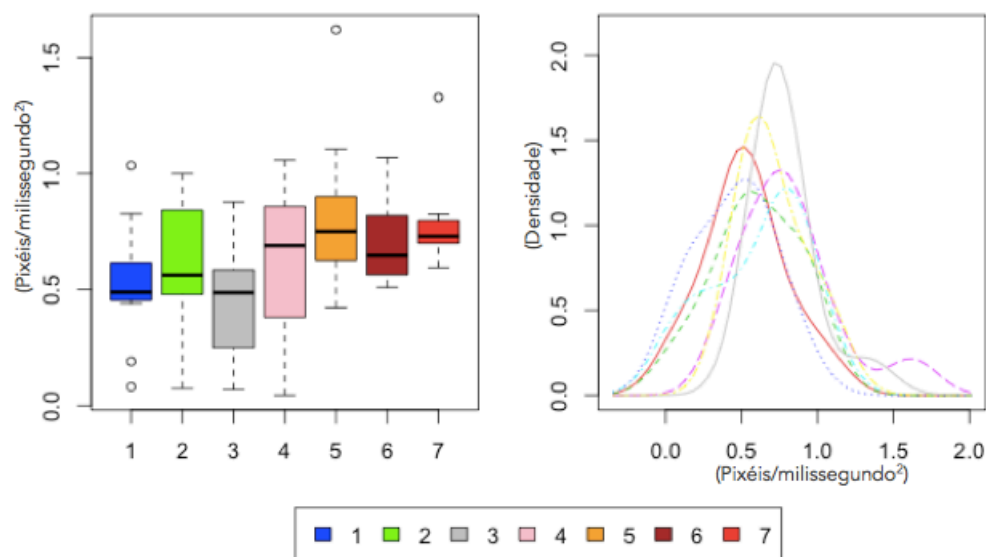


Figura 5.4: Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao MA

aumento da fadiga.

Tabela 5.7: Valores médios dos dados para o MV

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	0.48595	0.45726	0.90573	0.09049
Fatigado	0.7054	0.6909	1.6612	0.2209

5.3.3.2 Escala USAFSAM

Na tabela 5.8 e figura 5.6 podemos verificar o resultado da análise das distribuições para os 7 níveis de fadiga mental da escala USAFSAM relativamente à métrica *mouse velocity*. Verifica-se que o nível que possui a média e mediana mais elevado trata-se do nível 5, com um valor médio de 0.7949 e mediana 0.7315. Por outro lado verifica-se que o nível com um valor médio e mediana mais baixo trata-se que do nível 1. À semelhança da análise análoga para a aceleração do rato, verifica-se que não foi encontrada uma relação linear entre a velocidade do rato e os níveis.

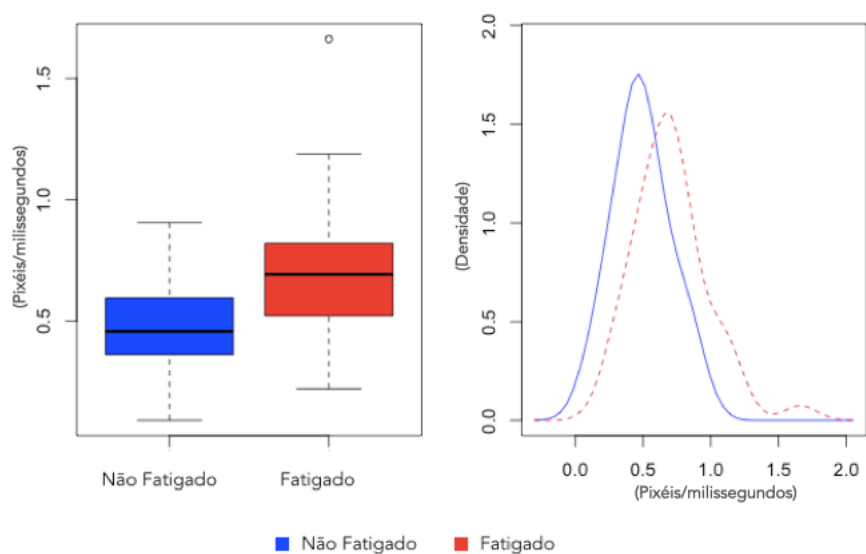


Figura 5.5: Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao MV

Tabela 5.8: Valores médios segundo a escala USAFSAM, relativas ao MV

Nível de fadiga mental	Média	Mediana	Máximo	Mínimo
1	0.4696	0.4464	0.8619	0.2279
2	0.5287	0.5321	0.8441	0.1769
3	0.47594	0.48656	0.90573	0.09049
4	0.6059	0.6441	1.0343	0.2209
5	0.7939	0.7315	1.6612	0.3675
6	0.6775	0.6099	1.1347	0.3997
7	0.7442	0.7271	1.1874	0.5642

5.3.4 Double Click Duration

A métrica DCD, tal como referido, mede a duração dos *clicks* duplos do rato. Esta medida pode relacionar-se com a reação e os tempo de resposta de um utilizador. Assim espera-se que nos casos de fadiga mental, esta métrica adquira valores mais elevados, uma vez que a reação é mais lenta, e os tempos de resposta mais elevados.

5.3.4.1 Ausência de fadiga vs estado de fadiga mental

A observação dos dados obtidos pelos utilizadores, como se pode verificar na figura 5.7 e tabela 5.9, indica que a média, mediana e valor máximo aumentou quando se compara os grupos "Não fatigado" e "Fatigado". Isto vai ao

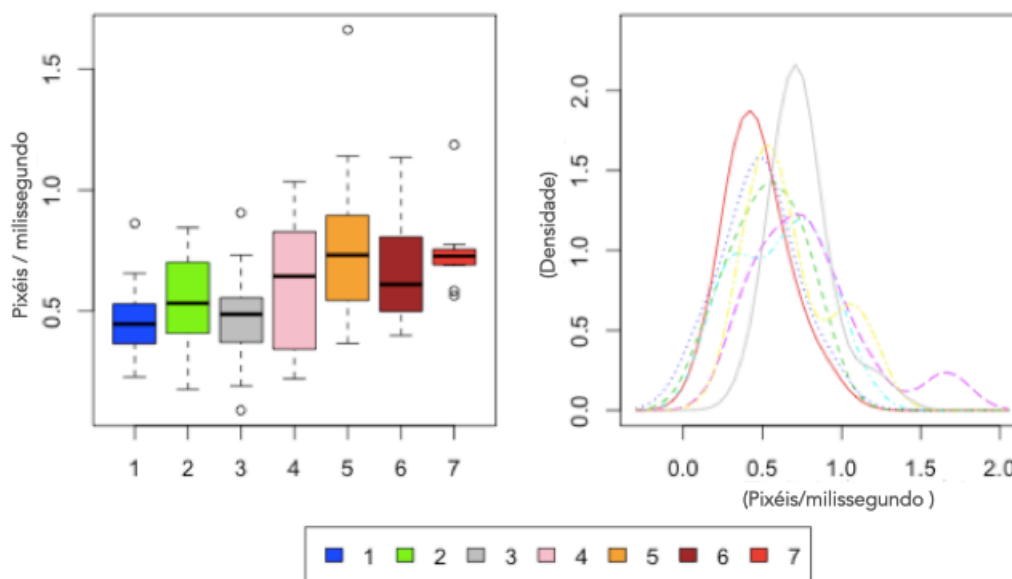


Figura 5.6: Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao MV

Tabela 5.9: Valores médios dos dados para o DCD

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	101.193	61.977 4	697.400	0.000
Fatigado	202.87	135.13	984.85	0.000

encontro daquilo que seria expectável.

5.3.5 Average Excess of Distance

Esta métrica mede excesso de distância que foi percorrido pelo rato entre dois pontos. Assim, é expectável que utilizadores fatigados tendam a percorrer distâncias maiores entre dois pontos, uma vez que a estes estados está associado uma menor precisão e falha de memória assim como menores níveis de concentração.

5.3.5.1 Ausência de fadiga vs estado de fadiga mental

Neste caso, na análise experimental, evidenciadas na tabela 5.10 e figura 5.8 verificou-se que um aumento do AED está associado a estados de fadiga mental. Tal facto é evidenciado pela média, em que o seu valor aumenta de 1.683 nos casos de ausência de fadiga, para 2.321 nos casos de fadiga mental. Na mediana, a diferença é ainda mais

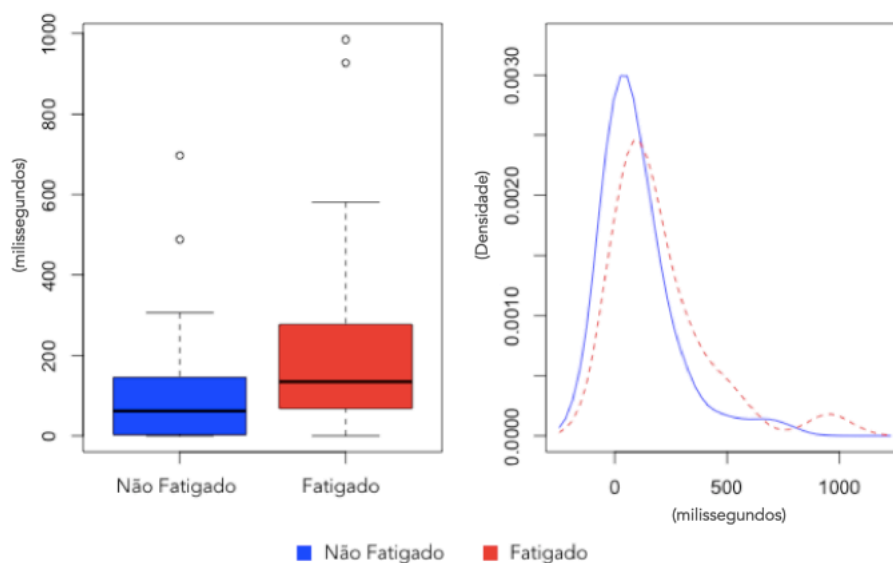


Figura 5.7: Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao DCD

expressiva, em que este caso aumenta de 1.633 para 1.998. Relativamente aos pontos extremos, no caso de fadiga mental estes também são superiores aos casos normais.

Tabela 5.10: Valores médios dos dados para o AED

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	1.683	1.633	2.048	0.000
Fatigado	2.321	1.998	10.678	1.300

5.3.5.2 Escala USAFSAM

Nesta análise verificou-se, tal como se pode observar na figura 5.9 e tabela 5.11, o nível de fadiga 5 trata-se do nível com a maior média e mediana, de 3.399 e 2.544 respetivamente. O nível com menor valor médio e mediano trata-se do nível 1 com 1.682 e 1.474. É possível ainda observar que o nível 1, 2 e 3 possuem valores médios muito semelhantes. À semelhança da análise efetuada para o MA e MV, não existe uma relação linear explícita entre o AED e o aumento do nível de cansaço.

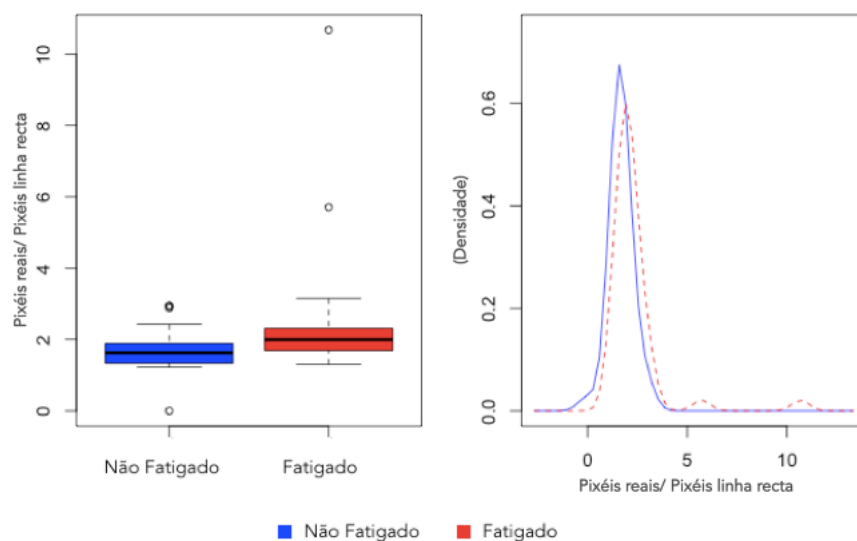


Figura 5.8: Histograma e diagrama de caixa para a distribuição "Não Fatigado" e "Fatigado", relativas ao AED

Tabela 5.11: Valores médios segundo a escala USAFSAM, relativas ao AED

Nível de fadiga mental	Média	Mediana	Máximo	Mínimo
1	1.682	1.474	2.941	1.223
2	1.686	1.681	2.048	1.357
3	1.684	1.688	2.889	0.000
4	2.079	1.996	2.644	1.606
5	3.399	2.544	10.677	1.364
6	1.810	1.845	2.204	1.300
7	1.994	1.797	3.132	1.419

5.3.6 Average Distance of the Mouse to the Straight Line

Nesta métrica, em que se avalia o comportamento do utilizador face à distância do rato até à linha reta. Neste sentido, espera-se a manifestação de uma menor precisão quando o utilizador está num estado de fadiga e por isso o aumento da distância média entre o caminho em linha reta o ponto onde o rato está posicionado.

5.3.6.1 Ausência de fadiga vs estado de fadiga mental

Através dos resultados obtidos, apresentados na figura 5.10 e tabela 5.12 pode-se verificar que tanto a média como a mediana sofrem um aumento quando são comparados utilizadores com estados mentais de fadiga, com utilizadores

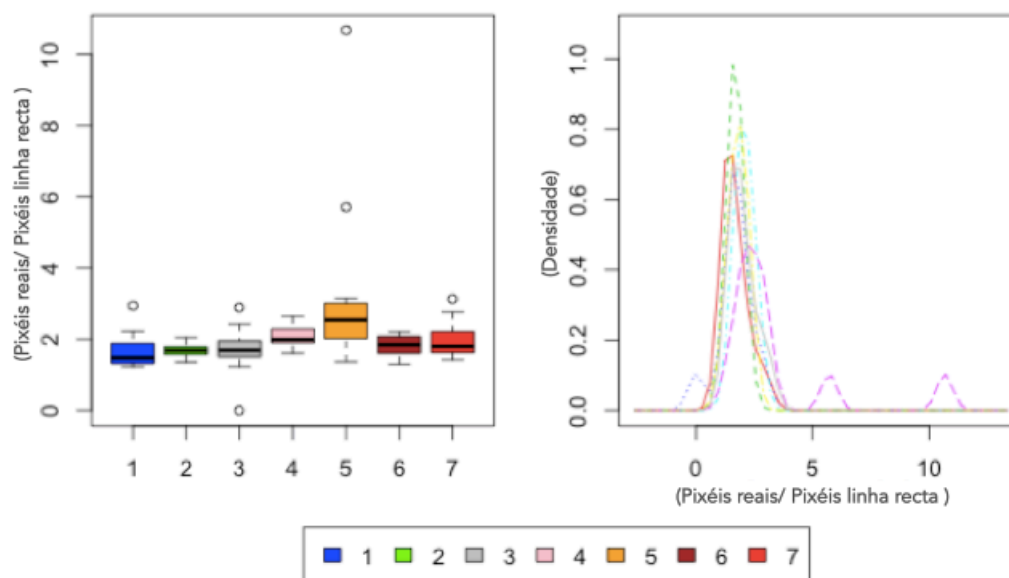


Figura 5.9: Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao AED

com estados de ausência de fadiga. No caso da mediana o aumento é mais acentuado, sendo que o seu valor aumenta de 121774 para 237387 pixels, e no caso da média de 175479 para 260747 pixels. Os extremos também aumentaram o seu valor.

Tabela 5.12: Valores médios dos dados obtidos para o ADMSL

Estado Mental	Média	Mediana	Máximo	Mínimo
Não Fatigado	121774	175479	121774	486174
Fatigado	237387	260747	237387	805350

5.3.6.2 Escala USAFSAM

Relativamente à análise com base nos níveis aferidos pela escala de USAFSAM, verificou-se que o nível com maior média e mediana trata-se do nível de fadiga 4, onde a média assume o valor de 319517 e a mediana o valor de 300674. Por outro lado verificou-se que o nível que toma a média mais baixa trata-se do nível 5, enquanto que a mediana mais baixa pertence ao nível 1. Verificou-se também que o ponto extremo mais elevado pertence ao 4, enquanto que o ponto extrema mais baixo pertence ao nível 5.

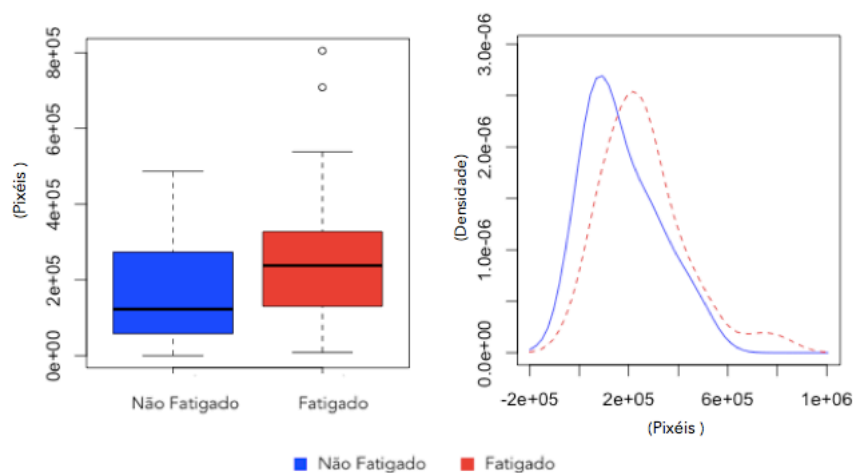


Figura 5.10: Histograma e diagrama de caixa para a distribuição "Não fatigado" e "Fatigado", relativas ao ADMSL

Novamente, pode-se concluir que não existe uma relação explícita entre o aumento do nível de fadiga com o aumento da ADMSL.

Tabela 5.13: Valores médios segundo a escala USAFSAM, relativas ao ADMSL

Nível de fadiga mental	Média	Mediana	Máximo	Mínimo
1	181344	104689	486174	43459
2	226892	274788	396164	83213
3	112336	57103	431595	486174
4	319517	300674	805350	107529
5	171226	164773	484294	8047
6	253431	236665	427090	128804
7	298814	253985	708700	60075

5.4 Conclusão

Após a finalização do estudo foi concluído que, para o *dataset* em questão a distinção entre um estado de fadiga mental e um estado de ausência de fadiga foi possível com base nas métricas KDT, MA, MV, DDC, AED e ADMSL, sendo que as métricas com mais influência foram pelo a MA e a MV. Relativamente à distinção entre os sete níveis de fadiga, segundo a escala USAFSAM o teste aplicado não permite aferir com exatidão que os sete níveis estão perfeitamente

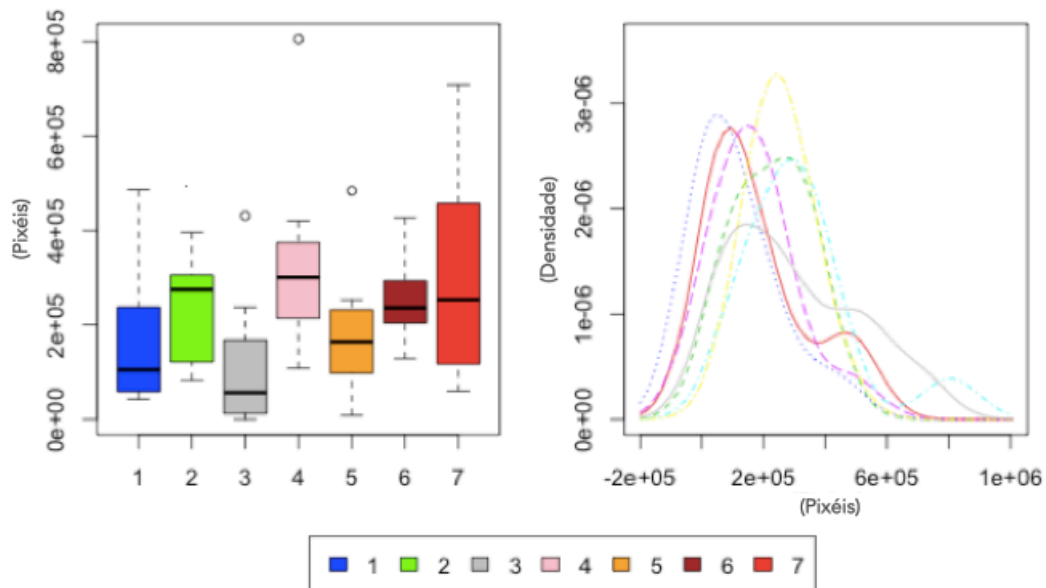


Figura 5.11: Histograma e diagrama de caixa para a distribuição segundo a escala USAFSAM, relativas ao ADMSL

distinguíveis, uma vez que apenas permite indicar que pelo menos duas amostras têm diferenças significativas. Apesar disso o teste tornou-se útil para despistar logo à partida quais as métricas que não teriam influência na distinção de qualquer nível. Desta forma, conclui-se que neste caso as métricas KDT, TBC, DDC, DMSL não teriam influência no estudo. Sendo assim, também foi concluído que as métricas consideradas relativas ao teclado não têm influência na distinção dos sete níveis de fadiga.

Com a análise efetuada foi também possível concluir que os comportamentos registados nem sempre foram de encontro ao esperado para cada métrica, o que indica que será erróneo a um pré-conceito acerca dos valores registados, uma vez que a fadiga é um fenómeno muito complexo e difícil de caraterizar.

Para além disto também foi possível verificar que não existe uma relação linear óbvia entre os níveis de fadiga segundo a escala USAFSAM e as métricas em estudo. Isto pode em parte ser devido ao facto que o *dataset* é proveniente de diferentes utilizadores, e não haver uma homogeneidade, assim como não se ter tido em conta a atividade que os utilizadores estavam a desempenhar.

Dados os resultados obtidos, pode-se ainda concluir que a utilização do rato em cenários de fadiga mental demonstrou um aumento das métricas relativas às distâncias percorridas, nomeadamente AED e ADMSL e uma aumento da velocidade e aceleração do rato, expressas através da MV e MA, o que poderá indicar que em cenários de fadiga mental, o utilizador tende a executar movimentos rápidos mas imprecisos. Relativamente às métricas relacionadas com o tempo KDT e DCD, verificou-se um aumento do tempo de resposta em casos de fadiga mental característicos de perfis destes estados.

Capítulo 6

Processo de aprendizagem

Uma vez finalizada a análise dos dados, iniciou-se a implementação de algoritmos não supervisionados de forma a detetar casos de fadiga mental. Com a aplicação dos algoritmos de aprendizagem não supervisionada pretendeu-se testar o agrupamento dos dados com base nas relações que os mesmos apresentam.

Tal como já foi referido, a deteção de fadiga mental é um processo muito complexo e subjetivo, por isso previamente foi definido uma estratégia de abordagem ao problema, com vista a obter os melhores resultados. Para tal, uma série de algoritmos não supervisionados foram escolhidos para o estudo de forma a compreender quais os que descreveriam o melhor comportamento. Posteriormente foram escolhidos os melhores modelos e implementados com o intuito de deteção de novos casos de fadiga.

6.1 Análise de agrupamentos

Numa primeira abordagem optou-se por uma análise de *clusters* com base nos estados mentais de ausência de fadiga e estados mentais de fadiga. Sendo assim, os dados foram agrupados em dois *clusters* segundo os algoritmos escolhidos. Após a obtenção dos grupos, foi importante uma análise detalhada dos mesmos de forma a identificar uma possível relação entre os agrupamentos obtidos e a classificação dos usuários. Esta relação foi possível, uma vez que o comportamento de cada estado mental, em função de cada métrica, já foi estudado na secção anterior.

Posteriormente, um outro *dataset* de igual dimensão ao anterior, foi criado apenas com instâncias pertencentes a um único utilizador. Como já mencionado anteriormente, os estados de fadiga mental são influenciados pelo perfil individual do utilizador. Sendo assim, o objetivo deste estudo passou por procurar uma melhoria da performance dos agrupamentos quando comparado com um modelo geral, onde são incluídos instâncias provenientes de vários utilizadores.

De seguida apresenta-se uma análise dos agrupamentos para os algoritmos não supervisionados selecionados.

6.1.1 K-means

Este algoritmo particional foi implementado com recurso à biblioteca *stats*¹ do R, mais especificamente à função *kmeans*. Optou-se pela inclusão deste algoritmo no estudo uma vez que se trata de um dos algoritmos mais simples, mas também um dos mais utilizados em aprendizagem não supervisionada.

6.1.1.1 Ausência de fadiga vs Estados de fadiga Mental

O algoritmo k-means foi aplicado a cada um dos *datasets* anteriormente referidos, com um $k=2$ e um $nstart=20$. Isto significa que serão obtido dois *clusters* (k), e serão testados vinte combinações diferentes de inicialização ($nstart$), sendo que posteriormente a função opta pela inicialização que demonstre menos variação dentro do agrupamento. O resultado apresenta-se na matriz de confusão representadas na tabela 6.1 que confronta o a classificação dos utilizadores, com o resultado do agrupamento obtido para cada um dos modelos testados.

Tabela 6.1: Matrizes de confusão para o algoritmo K-means, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		Modelo Geral	
		Normal	Fadiga
Grupo A		26	17
Grupo B		7	27

		Modelo Individual	
		Normal	Fadiga
Grupo A		25	8
Grupo B		8	31

Na tabela 6.2 pode-se observar a média das variáveis resultantes para os agrupamentos obtidos a partir do algoritmo k-means para o modelo individual e geral. Verifica-se no grupo A, à exceção do KDT para o modelo geral, todas as variáveis possuem menores valores quando comparado com o grupo B. Relativamente ao Modelo individual, a diferenças

¹<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/OOIndex.html>

entre as médias do grupo A e B não são tão significativas. Isto pode dever-se ao facto dos dados serem provenientes do mesmo indivíduo e consequentemente as métricas possuem menos variância, sendo por isso os valores mais próximos. No entanto pode-se verificar que em todas as métricas o grupo B, estas assumem uma média superior.

Tabela 6.2: Médias dos agrupamentos para o algoritmo K-means, na distinção entre ausência de fadiga e estados de fadiga mental

		KDT	MA	MV	DDC	AED	ADMSL
Modelo Geral	Grupo A	91.43	0.45912	0.44151	127.901	1.917	140473
	Grupo B	89.75	0.8723	0.8314	201.63	2.213	328201
Modelo Individual	Grupo A	87.41	0.62	0.59	104.045	1.776	202288
	Grupo B	93.88	0.6645	0.6359	215.42	2.313	243904

De forma a avaliar a qualidade dos agrupamentos encontrados foram adotadas índices de avaliação interna e externa de forma a compreender o quão úteis os agrupamentos encontrados são. Na tabela 6.3 encontra-se a avaliação interna dos *clusters* encontrados. Pode-se verificar que a qualidade dos agrupamentos encontrados não é elevada, sendo que os agrupamentos encontrados não estão claramente definidos, nem separados. No entanto, verifica-se que o modelo geral apresenta resultados ligeiramente superiores aos resultados do modelo individual. Apesar disso as diferenças entre os dois modelos não são significativas.

Tabela 6.3: Medidas de avaliação interna para o algoritmo k-means na comparação na distinção entre ausência de fadiga e estados de fadiga mental.

	índice de Dunn	índice de Davies-Bouldin	índice de silhouetta
Modelo Geral	0.08	1.67	0.26
Modelo Individual	0.11	1.96	0.20

Relativamente à avaliação externa dos *clusters*, verificou-se que o modelo individual apresenta um melhor agrupamento, aproximando-se mais dos resultados aferidos pelos utilizadores. Das 77 instâncias utilizadas em cada modelo, o modelo geral agrupou corretamente 53, enquanto que o modelo individual agrupou 56 instâncias corretamente. Desta forma, como se pode observar na tabela 6.4, o modelo individual obteve melhores resultados.

Tabela 6.4: Medidas de avaliação externa para o k-means

	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
Modelo Geral	68.8%	31.2%	0.794	0.614	0.692
Modelo Individual	72.7%	27.3%	0.795	0.795	0.795

6.1.2 PAM

O algoritmo Partition Around Medoids (PAM) foi aplicado aos *datasets* em questão recorrendo à livreria *cluster*² do R, mais especificamente à função *PAM*. Com a utilização deste algoritmo procurou-se melhorar os resultados relativamente ao K-means, uma vez que se trata de um algoritmo menos suscetível a *outliers*.

6.1.2.1 Ausência de fadiga vs Estados de fadiga Mental

Este algoritmo foi então aplicado aos dois *datasets* com um $k=2$, à semelhança do k-means. A matriz de confusão obtida da aplicação deste algoritmo apresenta-se na tabela 6.5. Como se pode verificar, no modelo geral, foi obtido um agrupamento com 30 instâncias (grupo A) e um outro grupo com 47 instâncias (grupo B). Relativamente ao modelo individual, o grupo A possui 27 instâncias, enquanto que o grupo B engloba 50 instâncias.

Tabela 6.5: Matrizes de confusão para o algoritmo PAM, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		Modelo Geral	
		Normal	Fadiga
Grupo A		22	8
Grupo B		11	36

		Modelo Individual	
		Normal	Fadiga
Grupo A		21	6
Grupo B		12	38

Analisando a média dos dados de cada um dos grupos, observa-se que à semelhança do k-means, o grupo B possui em ambos os modelos, uma média de todas as métricas mais elevada do que o grupo A, tal como demonstra a tabela 6.6.

Através da análise de avaliação dos índices internos, conclui-se mais uma vez que os agrupamentos encontrados não estão claramente definidos e separados, apresentado baixa performance para todos os índices em estudos. Quando se compara em concreto o Modelo Geral com o Modelo Individual verifica-se que as diferenças entre ambos são mínimas.

Relativamente à análise das medidas externas conclui-se que o Modelo individual se aproxima mais dos resultados

²<https://cran.r-project.org/web/packages/cluster/cluster.pdf>

Tabela 6.6: Médias dos agrupamentos para o algoritmo PAM, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		KDT	MA	MV	DDC	AED	ADMSL
Modelo Geral	Grupo A	79,60	0.405	0.396	98.93	1.8	118457
	Grupo B	97.76	0.7929	0.7525	199.7	2.203	290329
Modelo Individual	Grupo A	83.01	0.54777	0.53533	92.386	1.777	190832
	Grupo B	94.83	0.69224	0.6560	197.21	2.194	240935

Tabela 6.7: Medidas de avaliação interna para o algoritmo PAM

	índice de Dunn	índice de Davies-Bouldin	índice de silhouetta
Modelo Geral	0.069	1.67	0.21
Modelo Individual	0.066	2.08	0.16

aferidos pelos utilizadores, apesar da diferença ser muito pouco significativa. Como se pode verificar, que no modelo geral agrupou corretamente 58 instâncias, contra as 59 instâncias bem agrupadas do modelo individual.

Tabela 6.8: Medidas de avaliação externa para o PAM

	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
Modelo Geral	75.3%	24.7%	0.766	0.818	0.791
Modelo Individual	76.6%	23.4%	0.760	0.864	0.808

6.1.3 Método hierárquico aglomerativo

Este método hierárquico foi implementado com recurso à função *hclust*³ disponibilizada na biblioteca *stats* do R. No estudo efetuado foi utilizado como método a ligação simples. A principal razão para a inclusão deste algoritmo na presente dissertação foi a possibilidade de ser sensível a clusters com formas irregulares.

6.1.3.1 Ausência de fadiga vs Estados de fadiga Mental

Com o objetivo de obter dois agrupamentos, foi utilizada a função *cuttree*, que "corta" o dendrograma no número de grupos designado. Desta forma, a matriz de confusão obtida encontra-se evidenciada na tabela 6.9, onde se pode verificar que no modelo geral, foram obtidos dois grupos, denominados A e B, onde o grupo A possui 42 instâncias,

³<http://stat.ethz.ch/R-manual/R-devel/library/stats/html/hclust.html>

enquanto que o grupo B possui 35 instâncias. Relativamente ao modelo individual, o grupo A possui 28 instância, enquanto que o grupo B possui 49 instâncias.

Tabela 6.9: Matrizes de confusão para o algoritmo hierárquico aglomerativo, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		Modelo Geral	
		Normal	Fadiga
Grupo A		25	17
Grupo B		8	27

		Modelo Individual	
		Normal	Fadiga
Grupo A		17	11
Grupo B		16	33

Relativamente à média de cada um dos grupos encontrados, presente na tabela 6.10, verifica-se que no Modelo Geral, o agrupamento B possui uma média superior em todas as variáveis à exceção do DDC, onde o comportamento é inverso. Em relação ao Modelo Individual, verifica-se que não há uma tendência muito definida, sendo que os valores médios estão muito próximos. Apesar disso, verifica-se que o grupo A possui uma média mais elevada para as variáveis KDT, DDC, AED e ADMSL. Assim sendo, considerou-se que as características dos agrupamentos B em cada um dos agrupamentos mais semelhantes com os dados de utilizadores com fadiga mental.

Tabela 6.10: Medias dos agrupamentos para o algoritmo hierárquico aglomerativo, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		KDT	MA	MV	DDC	AED	ADMSL
Modelo Geral	Grupo A	94,93	0.452	0.438	186.40	1.931	153658
	Grupo B	85.59	0.8692	0.8245	129.32	2.188	307015
Modelo Individual	Grupo A	90.44	0.6366	0.6109	90.879	1.891	195370
	Grupo B	90.85	0.6282	0.5935	162.23	1,959	225094

Na tabela 6.11 encontra-se uma avaliação interna dos agrupamentos. Pela análise da tabela conclui-se mais uma vez que os índices revelam uma baixa qualidade de agrupamentos, sendo que os *clusters* não se encontram nem bem compactados, nem bem separados. Apesar disso, verifica-se que o Modelo Geral possui melhor resultados em todos os índices e por isso melhores agrupamentos.

Tabela 6.11: Medidas de avaliação interna para o algoritmo hierárquico aglomerativo

	índice de Dunn	índice de Davies-Bouldin	índice de silhouetta
Modelo Geral	0.092	1.73	0.23
Modelo Individual	0.078	4.512	0.016

Relativamente aos resultados obtidos da avaliação externa, verificou-se que contrariamente aos restantes algoritmos, o Modelo Geral obteve melhores resultados nas métricas em estudo quando comparado com o Modelo Individual. Sendo que, no Modelo Geral foram agrupadas corretamente 52 instâncias, enquanto que no Modelo Individual apenas foram agrupadas corretamente 50 instâncias.

Tabela 6.12: Medidas de avaliação externa para o método hierárquico aglomerativo

	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
Modelo Geral	67.53%	32.47%	0.771	0.613	0.683
Modelo Individual	64.93%	35.07%	0.763	0.750	0.709

6.1.4 Fuzzy c-means

O algoritmo fuzzy c-means foi implementado recorrendo à livreria *e1071*⁴, mais concretamente à função *cmeans*. A vantagem da inclusão deste algoritmo passa pela capacidade de lidar com *clusters* sobrepostos.

6.1.4.1 Ausência de fadiga vs Estados de fadiga Mental

De forma a obter dois agrupamentos foi utilizado um $k=2$. O resultado obtido deste algoritmo encontra-se na tabela de confusão 6.13 que confronta o resultado do agrupamento obtido com as classes aferidas pelos utilizadores. Pode-se verificar que do Modelo Geral resultaram dois grupos, um denominado Grupo A que engloba 43 das 77 instâncias iniciais, e um outro, o Grupo B que abrange 34 das instâncias iniciais. Relativamente ao Modelo Individual, verifica-se que o Grupo A inclui 39 instâncias, enquanto que o Grupo B abrange 38 instâncias.

Tabela 6.13: Matrizes de confusão para o algoritmo fuzzy c-means, na distinção entre estados de ausência de fadiga e estados de fadiga mental

		Modelo Geral	
		Normal	Fadiga
Grupo A		26	17
Grupo B		7	27

		Modelo Individual	
		Normal	Fadiga
Grupo A		25	14
Grupo B		8	30

Pela análise das médias de cada grupo, presente na tabela 6.14, podemos concluir que no caso do Modelo Geral, o valor de todas as métricas sofre um acréscimo quando comparado o Grupo A com o Grupo B. Apenas não se observa

⁴<https://cran.r-project.org/web/packages/e1071/e1071.pdf>

este comportamento para a variável KDT. Em relação ao Modelo Individual, observa-se que o Grupo B possui valores de todas as variáveis mais elevados, quando comparados com o Grupo A, apesar disso as diferenças entre as médias dos dois grupos é mais evidente no Modelo Geral do que no Modelo Individual.

Tabela 6.14: Médias dos agrupamentos para o algoritmo fuzzy c-means, na distinção entre ausência de fadiga e estados de fadiga mental

		KDT	MA	MV	DDC	AED	ADMSL
Modelo Geral	Grupo A	91.43	0.459	0.442	127.901	1.917	140473
	Grupo B	89.75	0.8723	0.8314	201.63	2.213	328201
Modelo Individual	Grupo A	88.21	0.6314	0.6023	140.04	1.765	213254
	Grupo B	93.22	0.652	0.6254	181,41	2,338	233744

Na avaliação interna deste agrupamento, foi considerado, adicionalmente, o índice silhueta de fuzzy, uma vez que o índice de silhueta convencional não considera a sobreposição que poderá existir entre os grupos. Assim, na tabela 6.15 podemos verificar os resultados da avaliação interna aplicada aos modelos. Verifica-se mais uma vez que os agrupamentos não são satisfatórios do ponto de vista estrutural. Apesar disso, observa-se que o índice de silhueta de fuzzy obteve um melhor resultado, quando comparado com o índice de silhueta comum, o que poderá indicar que os dois agrupamentos encontrados se encontram sobrepostos.

Tabela 6.15: Medidas de avaliação interna para o algoritmo fuzzy c-means

	índice de Dunn	índice de Davies-Bouldin	índice de silhouetta	índice de silhouetta de fuzzy
Modelo Geral	0.077	1.67	0.38	0.45
Modelo Individual	0.103	1.97	0.29	0.38

Tendo em consideração a avaliação externa pode-se verificar que o Modelo Individual obteve melhores resultados nas métricas em questão. Desta forma o Modelo Geral apresentou 53 instâncias bem agrupadas contra e 24 mal agrupadas, enquanto que o Modelo Individual agrupou 55 instâncias corretamente e 22 instâncias mal agrupadas. Os resultados dos índices de avaliação externa apresentam-se na tabela 6.16.

Tabela 6.16: Medidas de avaliação externa para o algoritmo fuzzy c-means

	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
Modelo Geral	63.53%	36.47%	0.794	0.614	0.692
Modelo Individual	71.42%	28.58%	0.789	0.682	0.731

6.2 Previsão de fadiga

Uma vez concluído o estudo de agrupamento dos dados, partiu-se para a previsão de novas instâncias de forma a compreender a eficácia dos modelos encontrados na detecção de fadiga mental. Para tal, foram selecionadas as situações que apresentaram um melhor agrupamento dos dados segundo as avaliações externas efetuadas. Assim, verificou-se que os algoritmos que apresentaram melhores performances trataram-se do K-means e PAM, na distinção entre estados de ausência de fadiga e estados de fadiga mental.

Na previsão das novas instâncias foram utilizadas 30 novas instâncias 15 de cada estado mental, para cada um dos modelos em questão.

6.2.1 k-means

Com o intuito de classificar as novas instâncias foi utilizada a menor distância entre o ponto médio encontrado durante a fase de agrupamento, e o novo ponto, usando para isso o calculo da distância euclidiana.

Desta forma um ponto pertence ao *cluster* cuja a distância euclidiana seja a menor entre esse mesmo ponto e o centro do *cluster*. Na tabela 6.17 e 6.18 pode-se observar o Modelo de Classificação Individual apresenta uma precisão mais elevada, sendo que conseguiu classificar corretamente 25 instâncias, enquanto que o Modelo de Classificação Geral classificou 22 instâncias corretamente. Isto traduziu-se em 83.3% de classificações corretas contra 73.3 % . Relativamente ao índice de concordância (Kappa s) verifica-se igualmente que o Modelo Individual obteve um melhor resultado. Observou-se igualmente, através do RMSE e MAE que o Modelo Individual também obteve menos erros de classificação.

Tabela 6.17: Matriz de confusão na previsão de novos casos de fadiga para o algoritmo k-means

		Modelo Geral	
		Normal	Fadiga
Grupo A		13	6
Grupo B		2	9

		Modelo Individual	
		Normal	Fadiga
Grupo A		13	3
Grupo B		2	12

Tabela 6.18: Índices de avaliação de previsão dos novos casos para o algoritmo k-means

	Classificação correta	Classificação Incorreta	RMSE	MAE	Kappa s
Modelo Geral	73.3%	26.7%	0.491	0.241	0.51
Modelo Individual	83.3%	16.7%	0.365	0.133	0.73

6.2.2 PAM

Nesta caso, para a classificação dos novos casos, foi utilizado o mesmo raciocínio que foi explicado para o k-means, no entanto não foi utilizado a média do *cluster* para o cálculo da distância, mas sim o ponto central, centroide. A tabela 6.19 e 6.20 o Modelo de Classificação Individual apresenta uma melhor performance na distinção entre estados mentais de fadiga e estados de ausência de fadiga. Observa-se que o Modelo geral obteve mais classificações corretas e menos classificações incorretas. Verifica-se, no entanto, que no geral os resultados não são tão satisfatórios quantos os obtidos pelo K-means, o que poderá indicar que a representação do centro do clusters com um ponto, ao contrário da utilização de uma média, poderá introduzir mais erros durante a previsão e detecção de fadiga em novos casos.

Tabela 6.19: Matriz de confusão na previsão de novos casos de fadiga para o algoritmo PAM

	Modelo Geral			Modelo Individual	
	Normal	Fadiga		Normal	Fadiga
Grupo A	8	5	Grupo A	10	1
Grupo B	7	10	Grupo B	5	14

Tabela 6.20: Índices de avaliação de previsão dos novos casos para o algoritmo PAM

	Classificação correta	Classificação Incorreta	RMSE	MAE	Kappa s
Modelo Geral	60.0%	40.0%	0.615	0.379	0.25
Modelo Individual	80.0%	20.0%	0.483	0.23	0.53

6.3 Discussão de resultados

Neste capítulo foram apresentados os resultados relativos ao agrupamento de casos de fadiga mental e posteriormente a previsão de novos casos. Relativamente à distinção entre casos de fadiga mental e estados de ausência de fadiga, pode-se concluir que apesar das métricas internas apresentarem resultados muito baixos, foi possível obter resultados razoáveis na distinção entre estes dois grupos, sendo por isso possível a sua distinção através de algoritmos não supervisionados. Devido à discrepância dos resultados entre as métricas internas e externas pode-se concluir, que neste caso, as métricas internas não são um bom índice de qualidade dos agrupamentos encontrados. Isto porque não têm em consideração o output esperado e apenas a estrutura interna dos agrupamentos.

Relativamente aos modelos em estudo : geral e individual verificou-se que à exceção do algoritmo hierárquico aglomerativo de ligação simples, todos os restantes apresentaram uma melhor avaliação externa para o modelo individual, tal encontra-se evidenciado na tabela 6.21 e tabela 6.22. Verificou-se, ainda, que os melhores resultados foram apresentados pelo PAM, onde foram agrupadas corretamente 76.6% no modelo individual e 72.7 % no modelo geral, seguido pelo k-means com 72.2% das instâncias bem agrupadas para o modelo individual. Sendo assim, conclui-se que a deteção de fadiga mental poderá passar pelo desenho de um modelo ajustado a cada individuo. Apesar disso, deve-se considerar que o estudo em causa apenas foi efetuado para um dos utilizadores, o que poderá levantar alguma reserva face aos resultados.

Tabela 6.21: Tabela comparativa para o modelo geral

Algoritmo	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
K-means	68.8%	31,2%	0.794	0.614	0.692
PAM	75.3%	24.7%	0.766	0.818	0.791
Fuzzy-c-means	65.53%	36.47%	0.794	0.614	0.692
Ligação-Simples	67.53%	32.47%	0.771	0.613	0.683

Tabela 6.22: Tabela comparativa para o modelo individual

Algoritmo	Instâncias bem agrupadas	Instâncias mal agrupadas	Precisão	Revocação	F1
K-means	72.7%	27.3%	0.795	0.795	0.795
PAM	76.6%	23.4%	0.760	0.864	0.808
Fuzzy c-means	71.42%	28.58%	0.789	0.682	0.731
Ligação-Simples	64.93%	35.07%	0.763	0.750	0.709

Tendo em consideração os resultados obtidos na previsão de novos casos de fadiga, conclui-se que o algoritmo que apresentou uma maior eficácia na previsão de novas instâncias foi k-means, tal como pode ser verificado na tabela 6.23, e que novamente o modelo individual apresentou melhores resultados destacando-se assim a importância do perfil do utilizador na deteção de fadiga.

Apesar dos resultados satisfatórios apresentados na distinção entre agrupamentos de casos de fadiga mental e casos de ausência de fadiga, o mesmo não foi conseguido na distinção entre os sete níveis de fadiga da escala USAFSAM. Uma vez que esta se trata de uma escala com muitos níveis, os métodos não supervisionados não se revelaram efi-

Tabela 6.23: Tabela resumo da previsão de novos casos de fadiga

		Classificação Correta	Classificação Incorreta
K-means	Modelo Geral	73.3%	26.7%
	Modelo Individual	83.3%	16.7%
PAM	Modelo Geral	60%	40%
	Modelo Individual	80%	20%

cazes em encontrar distinções entre os níveis. Por este motivo, apesar dos esforços foi impraticável esta distinção. Isto pode dever-se não só ao facto de que o aumento do números dos níveis estar inerente uma menor precisão na auto-avaliação dos indivíduos, mas também a uma sobreposição entre os dados dos diferentes níveis que poderá só ser resolvido recorrendo a aprendizagem supervisionada. Um exemplo do *clustering* obtido para o algoritmo k-means é apresentado em Apêndice A . Apesar do que já foi discutido, este resultado vai ao encontro da grande maioria dos estudos, que também apenas implementou uma deteção de fadiga com em dois níveis: ausência ou presença de fadiga.

Comparando com os resultados anteriores efetuados na área, considera-se que os resultados encontram-se enquadrados com o estado da arte relativo à deteção de fadiga recorrendo ao uso de teclado e rato. No entanto, nestes estudos existe uma deteção de fadiga enquadrada dentro de 5 e 7 níveis, o que no caso do presente estudo, como já referido, apenas foi possível a deteção entre dois níveis. Apesar disso deve-se considerar que estes estudos recorreram ao uso de aprendizagem supervisionada o que pode explicar o seu enquadramento dentro de uma escala de fadiga com mais níveis.

Capítulo 7

Conclusões e Trabalho Futuro

A fadiga mental apresenta-se como um conceito complexo, multifacetado e subjetivo, conhecido como uma sensação de cansaço e perda de performance após longos períodos de atividade cognitiva. De facto, esta problemática surge naturalmente no desempenho das tarefas diárias, sendo que muitas vezes é ignorada e olhada com alguma "leveza". Certo é que a fadiga mental afeta negativamente o nosso desempenho e produtividade, gerando situações de risco que podem estar na origem de graves falhas humanas. Assim, surge a necessidade da conceção de sistemas capazes de detetar e monitorizar estes estados mentais. Atualmente a deteção de fadiga é efetuada recorrendo a questionários o que torna a sua avaliação subjetiva e não permite uma monitorização continuada. Apesar do campo da viação já contar com uma vasta gama de soluções para a monitorização da problemática em causa, estas, na sua maioria, necessitam de algum tipo de hardware que se pode tornar incomodo ou dispendioso.

Os computadores são *devices* cada vez mais presentes no nosso quotidiano, raro é o local de trabalho, ou casa, onde não estejam presentes. Consequentemente o tempo dispensados na interação com os mesmos é cada vez maior. Esta interação está de tal forma desenvolvida que deu origem a um novo campo de estudo, a HCI. Duas das interfaces mais tradicionais de interação entre o homem e o computador trata-se do rato e do teclado. A interação baseada na utilização destas interfaces permite a recolha de métricas que já se revelaram úteis no passado na deteção de estados de fadiga mental através de métodos supervisionados. Apesar dos bons resultados alcançados, esta abordagem apresenta como principal desvantagem a necessidade da auto-avaliação por parte dos utilizadores. Assim, a presente monografia têm como proposta o estudo de um sistema não supervisionado de fadiga mental que permita a sua deteção.

7.1 Considerações sobre o trabalho

O sistema desenvolvido pela *startup Performetric* foi o ponto de partida para a presente monografia. Ao longo da mesma foi avaliado e estudado um conjunto de métricas de forma a compreender quais as que teriam relevância para a problemática em causa. Após a aplicação de testes não paramétricos como foi o caso do teste de Mann-Whitney para a distinção entre estados de fadiga e estados de ausência de fadiga e o teste de Kruskal-Wallis na distinção dos sete níveis de fadiga segundo a escala USAFSAM, foi possível retirar algumas conclusões acerca da forma como as métricas se relacionam com estes estados:

- Na distinção entre os estados de fadiga mental e estados de não fadiga as métricas TBC e DMSL não têm relação com estes estados;
- As métricas MA e MV são aquelas que possuem uma maior relação na distinção dos estados anteriormente referidos;
- Conseguiu-se estabelecer uma relação entre a diminuição de performance associadas à fadiga mental que se refletiram através do aumento das métricas DCD, AED, ADMSL e KDT. Apesar disso as métricas MV e MA revelaram um comportamento oposto ao que seria expectável. Assim, relativamente às métricas recolhidas pelo rato um aumento da AED e ADMSL acompanhado de um aumento do MV e MA, sugere que quando fatigados, os utilizadores apresentam comportamentos de manipulação do rato apesar de rápidos, imprecisos.
- Relativamente à relação entre as métricas em estudo e a sua relação com a escala USAFSAM, concluiu-se que estas não apresentam um comportamento linear, ou seja, um aumento de nível de fadiga não segue um comportamento linear de acordo com a métrica em estudo

Considerando a deteção de fadiga mental através de aprendizagem não supervisionada, conclui-se primeiramente que um modelo de deteção baseado nos sete níveis aferidos pela escala USAFSAM foi impraticável. Apesar disso foi possível uma deteção binária e a distinção entre estados de fadiga mental e de ausência de fadiga. Tendo em consideração os quatro algoritmos seleccionados, conclui-se que o algoritmo mais eficiente tratou-se do K-means tendo obtido

83,3% de novos casos bem classificados.

Por último, conclui-se, que o perfil individual de cada utilizador terá um enorme peso numa correta deteção de fadiga mental, sendo que os sistemas deverão ser o mais personalizados e individualizados possíveis.

7.2 Trabalho futuro

Embora se considere que foram atingidos grande parte dos objetivos da proposta dissertação, considera-se que existem melhorias que deveriam ser estudadas, e portanto algum trabalho futuro é deixado como consideração:

- Estudar de novas métricas e parâmetros relevantes para a problemática em causa. Também seria interessante a análise de cada *device* de *input* individualmente, e a comparação com o seu desempenho em conjunto;
- Alargar a amostra da população sobre o qual o estudo incide uma vez que esta pode ser considerada pequena;
- Analisar a manifestação de outras componentes muito importantes e que por vezes poderão confundir-se com fadiga mental, como é o caso do cansaço, sonolência, emoções.
- Testar outro tipo de algoritmos não supervisionados não explorados nesta dissertação, nomeadamente redes neurais não supervisionadas;
- Implementar um sistema distribuído baseado no modelo cliente-servidor de forma a que qualquer utilizador consiga fazer uso destas de um sistema não supervisionado de deteção automática de fadiga mental.

Bibliografia

- [1] Shyh-Yueh Cheng 1 e Hong-Te Hsu2. Mental Fatigue Measurement Using EEG. *Risk Management Trends*, 2011.
- [2] Hoda M. Abd-Elfattah, Faten H. Abdelazeim e Shorouk Elshennawy. Physical and cognitive consequences of fatigue: A review. *Journal of Advanced Research*, 6(3):351–358, 2015.
- [3] Periklis Andritsos. Data clustering techniques. *Toronto, University of Toronto, Dep. of Computer ...*, pages 1–34, 2002.
- [4] Colin M. Shapiro Azmeh Shahid , Kate Wilkinson, Shai Marcu. Visual Analogue Scale to Evaluate Fatigue Severity (VAS-F). *STOP, THAT and One Hundred Other Sleep Scales*, pages 399–402, 2011.
- [5] Sally Bailes, Eva Libman, Marc Baltzan, Rhonda Amsel, Ron Schondorf e Catherine S. Fichten. Brief and distinct empirical sleepiness and fatigue scales. *Journal of Psychosomatic Research*, 60(6):605–613, 2006.
- [6] Linsey Marinn Barker. *Measuring and modeling the effects of fatigue on performance: Specific application to the nursing profession*. PhD thesis, 2009.
- [7] Jacob Benesty, , Jingdong Chen, , Yiteng Huang, e Israel Cohen. Pearson Correlation Coefficient. In *Noise Reduction in Speech Processing*, pages 1–4. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [8] L.M. Bergasa, J. Nuevo, M.A. Sotelo, R. Barea e M.E. Lopez. vipO_Real-Time System for Monitoring Driver Vigilance_IEEE TRANSACTION. *IEEE Transactions on Intelligent Transportation Systems*, 7(1):63–77, 2006.
- [9] Ken Black. Business statistics contemporary decision making. *Journal of Chemical Information and Modeling*, 53(9):906, 2010.

- [10] Maarten A S Boksem, Theo F. Meijman e Monicque M. Lorist. Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research*, 25(1):107–116, 2005.
- [11] Maarten A.S. Boksem, Theo F. Meijman e Monicque M. Lorist. Mental fatigue, motivation and action monitoring. *Biological Psychology*, 72(2):123–132, 2006.
- [12] Maarten A.S. Boksem e Mattie Tops. Mental fatigue: Costs and benefits. *Brain Research Reviews*, 59(1):125–139, 2008.
- [13] P.A. Booth. Introduction to Human-Computer Interaction. *Manager*, pages 1–29, 1989.
- [14] Antônio de Pádua Braga, Andre Carvalho e Teresa Ludermir. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC, 2000.
- [15] Davide Carneiro, André Pimenta, Sérgio Gonçalves, José Neves e Paulo Novais. Monitoring and improving performance in human–computer interaction. *Concurrency and Computation: Practice and Experience*, 2015.
- [16] Abhijit Chaudhuri e Peter O Behan. Fatigue and basal ganglia. *Journal of the Neurological Sciences*, 179(1):34–42, 2000.
- [17] Abhijit Chaudhuri e Peter O Behan. Fatigue in neurological disorders chaudhuri.behan04. *The Lancet*, 363:1–11, 2004.
- [18] Antonio R Damasio. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, 1995.
- [19] Arturo De La Escalera, Marco Javier Flores e José María Armingol. Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions. *Eurasip Journal on Advances in Signal Processing*, 2010, 2010.
- [20] Alan Dix. Human computer interaction, foundations and new paradigms. *Journal of Visual Languages & Computing*, pages 1–13, 2016.
- [21] Alan Dix, Janet Finlay, Gregory Abowd e Russel Beale. *Human–Computer Interaction*, volume 48. Pearson, third edit edition, 2004.

- [22] N. Edenborough, R. Hammoud, A. Harbach, A. Ingold, B. Kisaćanin, P. Malawey, T. Newman, G. Scharenbroch, S. Skiver, M. Smith, A. Wilhelm, G. Witt, E. Yoder e H. Zhang. Driver State Monitor from DELPHI. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2(July):1206–1207, 2005.
- [23] F. Darby and C.Walls. *Stress and Fatigue Their impact on health and safety in the workplace*. Occupational Safety and Health Service of the Department of Labour, 1 edition, 1998.
- [24] Mehran Fazli e Masoud Shayesteh-azar. Correlation between the fatigue with gender , age and disease duration in multiple sclerosis patients. 2(11):206–209, 2013.
- [25] Richard Grace e Sonya Steward. Drowsy driver monitor and warning system. *International driving symposium on human factors in driver assessment, training and vehicle design*, pages 64–69, 2001.
- [26] Laurence Hartley, Tim Horberry, Nick Mabbott e G P Krueger. Review of Fatigue Detection and Prediction Technologies. *National Road Transport Commission*, (September):1–41, 2000.
- [27] Narelle L Haworth e Peter Vulcan. Testing of Commercially. (15), 1991.
- [28] Jesper F Hopstaken, Dimitri van der Linden, Arnold B Bakker e Michiel a J Kompier. The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological psychology*, 110:100–6, 2015.
- [29] Craig Jackson. The Chalder Fatigue Scale (CFQ 11). *Occupational Medicine*, 65(1):86, 2015.
- [30] Craig Jackson. The Chalder Fatigue Scale (CFQ 11). *Occupational Medicine*, 65(1):86, 2015.
- [31] a. K. Jain, M. N. Murty e P. J. Flynn. Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323, 1999.
- [32] E Joyce, S Blumenthal e S Wessely. Memory, attention, and executive function in chronic fatigue syndrome. *Journal of neurology, neurosurgery, and psychiatry*, (June):495–503, 1996.
- [33] Deepak Kapgate, Dhananjay Kalbande e Á Hybrid V-bcis Á. Advancements of Medical Electronics. pages 193–206, 2015.

- [34] Sinan Kaplan, Mehmet Amac Guvensan, Ali Gokhan Yavuz e Yasin Karalurt. Driver Behavior Analysis for Safe Driving: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3017–3032, 2015.
- [35] Fakhreddine Karray, Milad Alemzadeh, Jamil a Saleh e Mo N Arab. Human-Computer Interaction: Overview on State of the Art. *International Journal on Smart Sensing and Intelligent Systems*, 1(1):137–159, 2008.
- [36] Iftikhar Ahmed Khan e Nick Fine. Measuring Personality from Keyboard and Mouse Use. *Engineering and Technology*, (1984):1, 2008.
- [37] Preeti Khanna e M. Sasikumar. Recognising Emotions from Keyboard Stroke Pattern. *International Journal of Computer Applications*, 11(9):1–5, 2010.
- [38] Preeti Khanna, M Sasikumar e Kharghar Navi Mumbai. Recognising Emotions from Keyboard Stroke Pattern. *International Journal of Computer Applications*, 11(9):975–8887, 2010.
- [39] R. N. Khushaba, S. Kodagoda, S. Lal e G. Dissanayake. Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE Transactions on Biomedical Engineering*, 58(1):121–131, 2011.
- [40] Robert Langner, Michael B. Steinborn, Anjan Chatterjee, Walter Sturm e Klaus Willmes. Mental fatigue and temporal preparation in simple reaction-time performance. *Acta Psychologica*, 133(1):64–73, 2010.
- [41] Gang Li e Wan-Young Chung. Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. *Sensors (Basel, Switzerland)*, 13(12):16494–511, 2013.
- [42] Robert H Loblay. Chronic fatigue syndrome: Clinical Practice Guidelines. *Med J Aust*, 176(May):S17–S55, 2002.
- [43] Samuele M Marcora, Walter Staiano e Victoria Manning. Mental fatigue impairs physical performance in humans. *Journal of applied physiology (Bethesda, Md. : 1985)*, 106(3):857–864, 2009.
- [44] Emmanouil Michail, Athina Kokonozi, Ioanna Chouvarda e Nicos Maglaveras. EEG and HRV markers of sleepiness and loss of control during car driving. *The 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008:2566–2569, 2008.
- [45] Ajay Mittal, Kanika Kumar, Sarina Dhamija e Manvjeet Kaur. Head movement-based driver drowsiness detection : A review of state-of-art techniques. 2016.

- [46] Michael Negnevitsky. *Artificial Intelligence : A Guide to Intelligent Systems*. 2005.
- [47] Daniel Neu, Paul Linkowski e Olivier le Bon. Clinical complaints of daytime sleepiness and fatigue: how to distinguish and treat them, especially when they become 'excessive' or 'chronic'? *Acta neurologica Belgica*, 110(1):15–25, 2010.
- [48] Geri B. Neuberger. Measures of fatigue: The Fatigue Questionnaire, Fatigue Severity Scale, Multidimensional Assessment of Fatigue Scale, and Short Form-36 Vitality (Energy/Fatigue) Subscale of the Short Form Health Survey. *Arthritis & Rheumatism*, 49(S5):S175–S183, 2003.
- [49] André Pimenta. *Análise e Monitorização de Fadiga Mental*. PhD thesis, Universidade do Minho, 2013.
- [50] André Pimenta, Davide Carneiro, José Neves e Paulo Novais. A Neural Network to Classify Fatigue from Human-Computer Interaction. *Neurocomputing*, pages 413–426, 2016.
- [51] André Pimenta, Davide Carneiro, Paulo Novais e José Neves. Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8073 LNAI:222–231, 2013.
- [52] Andre Pimenta, Sergio Gonçalves, Davide Carneiro, Florentino Fde-riverola e Paulo Novais. Mental Workload Management as a Tool in e-Learning Scenarios. pages 25–32, 1994.
- [53] R Rojas. *Unsupervised Learning and Clustering Algorithms*. 1996.
- [54] Stuart J Russell e Peter Norvig. *Artificial Intelligence: A Modern Approach*, volume 9. 1995.
- [55] Arun Sahayadhas, Kenneth Sundaraj e Murugappan Murugappan. Detecting driver drowsiness based on sensors: A review. *Sensors (Switzerland)*, 12(12):16937–16953, 2012.
- [56] Sergio Salmeron-Majadas, Olga C. Santos e Jesus G. Boticario. An Evaluation of Mouse and Keyboard Interaction Indicators towards Non-intrusive and Low Cost Affective Modeling in an Educational Context. *Procedia Computer Science*, 35:691–700, 2014.

- [57] Sw Samn e Lp Perelli. Estimating Aircrew Fatigue: A Technique with Application to Airlift Operations. page 26, 1982.
- [58] Kota Sampei, Miho Ogawa, Carlos Torres, Munehiko Sato e Norihisa Miki. Mental Fatigue Monitoring Using a Wearable Transparent Eye Detection System. *Micromachines*, 7(2):20, 2016.
- [59] Mandalapu Saradadevi. Driver Fatigue Detection Using Mouth and Yawning Analysis. *International Journal of Computer Science and Network Secuirity*, 8(6):183–188, 2008.
- [60] SeeingMachines. faceLAB 5 - Eyetracking for Research. 2009.
- [61] Helen Sharp, Yvonne Rogers e Jenny Preece. Interaction design: beyond human-computer interaction. *Book*, 11:773, 2007.
- [62] Jianhua Shen, Joseph Barbera e Colin M. Shapiro. Distinguishing sleepiness and fatigue: Focus on definition and measurement. *Sleep Medicine Reviews*, 10(1):63–76, 2006.
- [63] Weijie SHEN; Haixin SUN; En CHENG; Qingkun ZHU; Qihu Li; Weijun SHEN. Effective Driver Fatigue Monitoring through Pupil Detection and Yawing Analysis in Low Light Level Environments. *SOURCEInternational Journal of Digital Content Technology & its Applic*, 6(17):372, 2012.
- [64] Karin C H J Smolders e Yvonne A W de Kort. Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal. *Journal of Environmental Psychology*, 39:77–91, 2014.
- [65] Mark A Staal. Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework. *National Aeronautics and Space Administration*, (August):168, 2004.
- [66] R Suganya e R Shanthi. Fuzzy C-Means Algorithm-A Review. *International Journal of Scientific and Research Publications*, 2(11):2250–3153, 2012.
- [67] Kazem Taghva e Rushikesh Veni. Effects of similarity metrics on document clustering. *ITNG2010 - 7th International Conference on Information Technology: New Generations*, pages 222–226, 2010.
- [68] Pang-Ning Tan, Michael Steinbach e Vipin Kumar. Chap 8 : Cluster Analysis: Basic Concepts and Algorithms. In *Introduction to Data Mining*, page Chapter 8. 2005.

- [69] Leonard J Trejo, Rebekah Kochavi, Karla Kubitz, Leslie D Montgomery, Roman Rosipal e Bryan Matthews. EEG-based Estimation of Cognitive Fatigue. *Psychological Assessment*, 5797:1–11, 2005.
- [70] Alexandros Tsoulouhas, Georgios; Georgiou, Dimitrios; Karakos. Detection of Learner's Affective State Based on Mouse Movements. *Journal of Computing; Nov2011, Vol. 3 Issue 11, p9*, 2011.
- [71] Rohit Tyagi, Kaiquan Shen, Shiyun Shao e Xiaoping Li. A novel auditory working-memory vigilance task for mental fatigue assessment. *Safety Science*, 47(7):967–972, 2009.
- [72] Anne Underwood. *The quest for artificial blood.*, volume 139. 2002.
- [73] Dimitri van der Linden, Michael Frese e Theo F Meijman. Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psychologica*, 113(1):45–65, 2003.
- [74] Esra Vural, Mujdat Cetin, Aytul Ercil, Gwen Littlewort, Marian Bartlett e Javier Movellan. Drowsy Driver Detection Through Facial Movement Analysis. *Human-Computer Interaction*, 4796:7–18, 2007.
- [75] Ming Cong Weng, Chia Tseng Chen e Hsiang Chun Kao. Remote surveillance system for driver drowsiness in real-time using low-cost embedded platform. *Proceedings of the 2008 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2008*, pages 288–292, 2008.
- [76] Shi Wenwen, Jiang Fuchuan, Zheng Qiang e Cui Jingjing. Analysis and Control of Human Error. *Procedia Engineering*, 26:2126–2132, 2011.
- [77] Mervyn V M Yeo, Xiaoping Li, Kaiquan Shen e Einar P V Wilder-Smith. Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science*, 47(1):115–124, 2009.
- [78] Hong Yu, LC Shi e Bao-Liang Lu. Vigilance Estimation Based on EEG Signals. ... of IEEE/ICME International Conference on ..., 2007.

Apêndices

Apêndice A

Clustering segundo a escala USAFSAM

Tabela A.1: Tabela k-means segundo a escala USAFSAM

Modelo Geral								Modelo Individual							
	1	2	3	4	5	6	7		1	2	3	4	5	5	7
A	1	2	3	2	0	0	0	A	1	0	0	5	2	5	2
B	0	0	0	0	1	0	0	B	0	1	2	0	0	0	2
C	1	1	0	2	1	1	3	C	6	5	3	3	5	0	0
D	1	1	4	0	1	1	4	D	0	0	0	1	0	0	0
E	6	1	2	2	3	1	1	E	2	2	1	0	0	0	2
F	1	3	1	2	5	2	1	F	1	0	3	0	3	1	0
G	1	3	1	3	0	6	3	G	1	3	2	2	2	5	5